UNIVERSITEIT
STELLENBOSCH
UNIVERSITY

Antitrust market definition using statistical learning techniques and consumer characteristics

Willem H. Boshoff

Working Paper 224

June 2011

Antitrust market definition using statistical learning techniques and consumer characteristics

Willem H. Boshoff[*]

June 22, 2011

Abstract

Market definition is the first step in an antitrust case and relies on empirical evidence of substitution patterns. Cross-price elasticity estimates are preferred evidence for studying substitution patterns, due to advances in IO econometric modelling. However, the data and time requirements of these models weigh against their universal adoption for market definition purposes. These practical constraints – and the need for a greater variety of evidence – lead practitioners to rely on a larger set of less sophisticated tools for market definition. The paper proposes an addition to the existing toolkit, namely an analysis of consumer characteristics for market definition purposes. The paper shows how cluster analysis can be used to identify meaningful groups of substitutes on the basis of homogeneity of their consumer profiles. Cluster analysis enforces consistency, while recent bootstrap techniques ensure robust conclusions. To illustrate the tool, the paper relies on data from a recently concluded radio merger in South Africa.

JEL codes L4 L1 C1 D1

Keywords market definition substitutes media demography clusters bootstrap

1 Introduction

In most jurisdictions market definition is the first step in a competition case and relies on empirical evidence of substitution patterns. Cross-price elasticity estimates are preferred forms of evidence for studying substitution patterns, due to advances in econometric modelling in the field of industrial organisation (IO) over the past two decades. However, onerous data and time requirements of econometric models in IO weigh against their universal adoption for market definition purposes. These practical constraints, and the fact that the models embody specific assumptions concerning firm behaviour, support the continued use of a larger set of less sophisticated tools for market definition. While less sophisticated tools individually rely only on a subset of information, the combination of different tools significantly expands the available information set and broadens the basis for inferences. This study proposes a complement to the existing set of tools, namely an analysis of consumer characteristics for market definition purposes. Specifically, the study shows how cluster analysis (a type of statistical learning technique) can be used to identify meaningful groups of substitutes on the basis of homogeneity of their consumer profiles. Cluster analysis enforces consistency, while recent advances now allow statistical inference to ensure robust conclusions. The relatively easy implementation and graphical output which assists interpretation further recommends cluster analysis as a tool for market definition.

Before presenting the technique, the study considers briefly the rationale for market definition in modern competition policy and the extent to which a study of consumer characteristics is consistent with the concept of an antitrust market. The study then discusses hierarchical cluster analysis, illustrating its application to a market definition problem in a recent radio merger case in South Africa.

2 Consumer characteristics and market definition

During market definition, the analyst identifies potential substitutes for the product sold by the firm(s) being investigated. The analyst then ranks these substitutes in terms of their degree of substitutability and includes higher-ranked substitutes in the relevant market based on a selection criterion. One popular criterion is offered by the hypothetical monopolist test, which is a thought experiment in which the sensitivity of buyers to a small but significant non-transitory increase in price (SSNIP) is used to define the market. In particular (United States Department of Justice and Federal Trade Commission 1992):

"A market is defined as a product or group of products and a geographic area in which it is produced or sold such that a hypothetical profit-maximizing firm, not subject to price regulation, that was the only present and future producer or seller of those products in that area likely would impose at least a 'small but significant non-transitory' increase in price, assuming the terms of sale of all other products are held constant. A relevant market is a group of products and a geographic area that is no bigger than necessary to satisfy this test".

The SSNIP thought experiment, as codified in US merger guidelines, is therefore explicitly focused on consumer behaviour. This focus on consumer behaviour was retained in the 2010 revisions of the merger guidelines .(United States Department of Justice and Federal Trade Commission 2010: 7):

"Market definition focuses solely on demand substitution factors, i.e., on customers' ability and willingness to substitute away from one product to another in response to a price increase or a corresponding non-price change such as a reduction in product quality or service".

Of course, supply-side features also matter for market definition, although these features are frequently considered in the assessment of market power rather than the delineation of the relevant market. A focus on consumer behaviour is also found in other competition jurisdictions, including Europe and developing countries such as South Africa (Motta 2006; Davis and Garcés 2010).

The application of the SSNIP test is straightforward in homogenous goods markets, but becomes quite difficult for differentiated goods. In differentiated goods markets it is preferable to rely on a broad set of information on consumer substitution patterns when defining the relevant market and the competition literature has suggested a range of empirical tools for market definition purposes. Consumer surveys providing diversion ratios (Katz and Shapiro 2003; O'Brien and Wickelgren 2003; Farrell and Shapiro 2010) and econometric models providing price elasticity estimates (Froeb and Werden 1991; Ivaldi and Verboven 2005) are the most direct tools for market definition, as they are directly relevant to the SSNIP question. Such direct forms of evidence on consumer behaviour are not always feasible to obtain, especially in developing market contexts. Even when data is available, time constraints may not permit sophisticated econometric modelling (Coate and Fischer 2008). Furthermore, beyond time and data constraints, competition authorities do not necessarily weigh any particular piece of quantitative evidence heavier than other forms of evidence, requiring direct evidence for market definition purposes to be corroborated by other evidence .(Bishop and Walker 2002). The practical problems of time, availability and information from different types of evidence motivate further development of the set of indirect tools for market delineation. It is against this background that the study develops an additional tool for market definition based on a statistical analysis of heterogeneity in consumer characteristics.

Beyond practical motivations, there are also deeper theoretical reasons for considering additional tools for market definition. The increasing emphasis on an effects-based approach signals a move

away from a structuralist approach to competition, which views competition as an outcome dictated by market structure. Empirical IO models continue to model structural market outcomes, which may not always sufficiently describe the extent of competition (Manne and Wright 2009). Therefore, it may be useful to complement the mainstream IO models with approaches that are less structuralist in nature.

These theoretical and practical challenges justify the use of complementary tools for market definition that view competition more as a rivalrous process and less as a structural characteristic of markets (see early critiques by Hayek (1946 [1984]) and McNulty (1967)). This study suggests one such a complementary tool, namely the study of heterogeneity in consumer characteristics. The competition policy literature suggests that, in addition to price evidence or product similarity, it is useful to consider the demographic profile of consumers of differentiated products. Coate and Fischer (2008: 33) note, for example, how the substitutability of differentiated products depend on the class of consumer: "two products that appear relatively similar to one customer may be highly differentiated to another". The use of consumer characteristics in studying substitution is not new and the recent IO literature has incorporated consumer characteristics to help improve elasticity estimates in empirical IO models. Before the 1980s, IO models made little allowance for consumer heterogeneity, focusing on representative agent models. Consumer characteristics became more important with the emergence of discrete choice models of differentiated products, which accounted for products as bundles of characteristics (McFadden 1981; Berry 1994), and built on the hedonic models developed earlier (Lancaster 1966; 1971; Gorman 1980). Specifically, Berry, Levinsohn and Pakes (1995) (BLP), Nevo (2001) and others modelled the interaction between product characteristics and consumer characteristics. In the earlier discrete choice models of product characteristics, substitution was modelled for average individuals with identical preferences. In contrast the newer models see buyers of a particular product as those consumers who value the bundle of characteristics of the particular product. When price increases, consumers switch to other products with similar characteristics. These models therefore allow "estimated substitution patterns and (thus) welfare to directly reflect demographic-driven differences in tastes for observed characteristics" (Petrin 2002).

While individual consumer characteristics data may be useful for market definition purposes (Berry, Levinsohn et al. 2004), data constraints may limit their use in competition investigations. However, while consumer-level data significantly improve estimates because of their level of disaggregation, the recent IO literature suggests that less disaggregated data on consumer characteristics can still offer significant insights into substitution. Petrin (2002), for example, incorporates data on the average demographic profile of consumers, calculated from market- rather than individual-level data, to arrive at improved elasticity estimates. Aggregate consumer characteristics therefore assist in elasticity estimation and, hence, the study of substitutability. Therefore, one may argue that a study of aggregate consumer characteristics can be useful for market definition.

A study of heterogeneity in consumer characteristics can be particularly useful for market definition in two types of competition investigation. Firstly, heterogeneity in consumer characteristics can assist in market definition under conditions where systematic price data is not available. This problem is particularly acute in media markets, such as radio stations and newspapers, where consumers do not pay in monetary terms for the product. In these cases, the absence of price data may prevent elasticity estimates or, even, the use of less sophisticated econometric tools such as price tests. Frequently, practitioners then revert to qualitative analysis. However, it is still possible to employ a systematic tool for market definition, if rich consumer characteristics data are available.

Secondly, heterogeneity in consumer characteristics can assist in market definition if it can be established, *a priori*, that the product under investigation is closely targeted at a specific consumer group. For example, the marketing literature shows that information and entertainment goods, say particular satellite TV music stations, are highly targeted goods aimed at very specific consumer groups. Therefore, differences in consumer characteristics closely approximate differences in product characteristics. While an assessment of "product" characteristics does not necessarily settle the market definition question, it provides a useful ranking of products that can assist in inferring the

likely response to a SSNIP.

More generally, even in competition investigations where these two conditions do not apply, a study of the heterogeneity of consumer characteristics remains a useful additional tool that provides an alternative perspective on the market and the likely response to a SSNIP. Nevertheless, the empirical application later in this chapter involves a particularly relevant case, as we focus on a media market where consumers do not pay in monetary terms and where the product is particularly targeted at a specific consumer group. Before considering this application, the following section discusses the tool called cluster analysis, which is used to analyse heterogeneity in consumer characteristics data on a statistical basis.

3 Cluster analysis

A study of heterogeneity in consumer characteristics assists in market definition because it provides estimates of price elasticity. Consider two firms, selling products a and b, proposing to merge. To define the relevant market(s), the competition analyst considers all candidate products, including a and b. Ideally, the analyst would study a demand system for the n candidate products and an accompanying $n \times n$ matrix of price elasticities, where the diagonal contains the n own-price elasticities and the off-diagonal elements the various cross-price elasticities. The off-diagonal elements would then be useful to assess the extent of substitution among a, b and the other $n - 2$ products and, hence, the relevant market(s). In the absence of formal elasticity estimates, the analyst may consider the consumer profiles for products a and b, captured by vectors c_a and c_b. The similarity of the consumer characteristics for the two products are then captured by the distance between c_a and c_b. This study argues that the similarity of the consumer characteristics, as measured by the distance, approximates the cross-price elasticity between products a and b. When the distance between c_a and c_b is small, the cross-price elasticity between the two products is high. Vice versa, the larger the distance between the characteristics vectors, the lower the cross-price elasticity. Therefore, if products are in the same market, their respective *consumer characteristics vectors will be similar and therefore tend to cluster together*, while those products that are outside of the market will not form part of this cluster. The statistical learning literature offers a number of data-based techniques to identify formally such clusters in data. Consequently, statistical learning techniques, as applied to consumer characteristics data, offer an additional tool for implementing the SSNIP test. The study focuses on one such a technique, hierarchical cluster analysis, which is discussed in the following section.

Cluster analysis is a well-known technique in the statistical learning literature and statisticians frequently apply cluster analysis to group multivariate observations (Lorr 1983; Kaufman and Rousseeuw 1990). A cluster is defined as a group of observations, where the observations are more similar to one another than they are to observations outside of the group (Everitt 1993; Johnson and Wichern 2002). Grouping is therefore never "objective" or "natural", but the result of a particular similarity concept employed; cluster analysis is not "a graphical summary of the data itself [but] a description of the results of the algorithms" (Hastie, Tibshirani et al. 2001: 475). In particular, cluster analysis depends on two similarity measures chosen by the analyst: firstly, the distance metric (a measure of similarity between observations) and, secondly, the linkage method (a measure of similarity between clusters).

3.1 Distance

The choice of distance metric is determined by the analyst's tolerance for dissimilarity. In market definition, the analyst wants to group product A (or, if it is a merger between two firms, product A and B) with competitors of the product(s) that satisfy the SSNIP test and seeks to avoid defining the market overly narrowly or overly broadly. This, in turn, requires a distance metric that does not overstate the differences between substitutes.

Note that cluster analysis for market definition purposes requires actual distances. In many competition cases, consumer surveys are conducted where participants are asked to judge the degree of substitutability among different products. These subjective "distance" measures are not directly useful for cluster analysis, as they frequently violate the triangle inequality[1] (Hastie, Tibshirani et al. 2001: 455). Cluster analysis applied to subjective distances will be misleading and the data must be transformed to dissimilarities for cluster analysis purposes. The transformed distance matrix obtained from this data may also be non-symmetric and will have to be transformed to symmetric form.

A distance metric should weigh appropriately the contributions of the individual variables in a multivariate observation, as some variables play a larger role in clustering than others. Euclidean distance is a squared distance and assigns greater weight to large differences than to smaller ones, so that variables with larger orders of variation will dominate more stable variables in the distance metric:

$$d(y_{i_1}, y_{i_2}) = [(y_{i_1} - y_{i_2})'(y_{i_1} - y_{i_2})]^{\frac{1}{2}} \tag{1}$$

where y_{t_1} and y_{t_2} have been mean-centred.

The emphasis on larger orders of variation is useful for market definition purposes, as the competition analyst may want to include products in the market that are slightly different, but may be much more wary to include them when they differ markedly. One could replace the squared distance with absolute deviation distance, but this will reduce the emphasis on larger differences and the study uses squared distance.

Two further challenges facing distance calculations are, firstly, differences in the scaling of different variables and, secondly, correlation among the different variables. The chi-squared distance offers a way of dealing with the scaling problem, by dividing data for each variable with the sample variance:

$$d(y_{i_1}, y_{i_2}) = [(y_{i_1} - y_{i_2})' A^{-1} (y_{i_1} - y_{i_2})]^{\frac{1}{2}} \tag{2}$$

where A is a diagonal matrix, containing sample variances for the different variables.

In addition, correlation among the variables may bias the Euclidean and chi-squared distance metrics downwards. The statistical distance metric addresses the correlation and scaling problem simultaneously by weighing the different variables with their respective sample variances and transforming the system of correlated coordinates to a system of uncorrelated coordinates:

$$d(y_{i_1}, y_{i_2}) - [(y_{i_1} - y_{i_2})' S^{-1} (y_{i_1} - y_{i_2})]^{\frac{1}{2}} \tag{3}$$

where S is the sample covariance matrix.

The study uses statistical distance in the empirical application.

3.2   Linkage method

The distance metric is a measure of the similarity of observations. Cluster analysis also requires a measure of similarity of clusters, known as the linkage method. The linkage method guides the decision concerning which lower-level clusters to merge: in each round, the linkage method chooses the two nodes with the smallest dissimilarity for merger into a single cluster. Cluster linkage methods differ according to size bias: some linkage methods have a higher tolerance level for dissimilarity within merged clusters and will tend to suggest larger clusters.

One should conduct a variety of cluster analyses to check the sensitivity of output for a particular linkage method, although the centroid and median methods tend to produce "inversion" (Morgan and Ray 1995). As discussed earlier, clustering starts with grouping the more similar items, so that successively larger clusters contain together less similar items. Therefore, the overall variance within clusters should increase as we move up the tree. Inversion refers to the violation of this property,

---

[1]The triangle inequality requires that $d(y_i, y_j) \leqslant d(y_i, y_k) + d(y_j, y_k)$ for all $k = 1, ..., N$.

where a cluster at a particular height has *lower* variance than its two constituent clusters. While inversion is not a systematic feature of the centroid and median linkage methods, Morgan and Ray (1995) show that these methods are more prone to produce inversion for particular types of data and is best avoided. The problem of inversion raises the more general question: which linkage method is best? The statistical literature suggests a number of linkage methods of which the single, average and complete linkage methods are the most prominent and widely used. Each of these methods suggests a particular parameterisation of the analyst's loss function.

The single linkage method is also known as the nearest neighbour linkage method (Johnson 1967). The method only fuses two lower-level clusters into a new cluster if the nearest members in the two groups are close enough (Everitt 1993). In a market definition sense, two hypothetical submarkets are fused into a larger submarket if the consumer profile of their nearest members is very similar. The problem with the single linkage method is that it is biased towards "chaining": the merger of two clusters only requires that one of the observations in each of the clusters is sufficiently close to one another. It will therefore tend to produce clusters with a large diameter (where diameter refers to the largest distance among the members of a particular cluster) (Johnson and Wichern 2002). This violates the compactness criterion for a cluster; under the single linkage method observations within clusters may differ quite substantially. The complete linkage method, on the other hand, is biased towards small clusters: clusters only merge if the pair of most remote observations in the two clusters is similar (Everitt 1993). The problem with this method is that it violates closeness, as some observations within the cluster are perhaps closer to observations outside of the cluster. In comparing the single and complete linkage methods, the question is clearly how important violations of "compactness" and of "closeness" are to the decision maker (Johnson 1967; Lorr 1983; Hastie, Tibshirani et al. 2001). Closeness is important to a competition analyst, as it ensures that one does not include competitors in the market when there are closer competitors outside of the market. Compactness may also be important, as the analyst does not want a market that includes too many irrelevant competitors.

The average linkage method strikes a balance between compactness and closeness. This method offers a compromise between the single and complete linkage methods and defines the similarity of clusters as the average distance between all pairs of observations in the two groups (Sokal and Michener 1958; Lance and Williams 1967). The average linkage method is sensitive to "the numerical scale on which the observation dissimilarities ... are measured" (Hastie, Tibshirani et al. 2001: 477): even if a strictly increasing transformation is applied, so that relative ordering of distances is preserved, the answers may change for different distance metrics. The single and complete linkage methods only require an ordering of the distances and are invariant to strictly increasing transformations (Hastie, Tibshirani et al. 2001). However, Kelly and Rice (1990) argue that the average linkage method produces clusters with a superior probability interpretation. Cluster analysis can be seen as attempts at identifying the local modes in a joint probability distribution and Kelly and Rice (1990) show that only clusters produced by average linkage methods bear any relation to characteristics of the joint probability distribution.

There are conditions under which different linkage methods generate quite different results (excluding those methods that suffer from inversion). The consumer characteristics data in the empirical application discussed later is one such an example, with the average linkage method results differing substantially from the single and complete linkage results. This is usually the case when there is not a strong clustering tendency in the data as a whole. It is therefore prudent to consider all three linkage methods described above when conducting cluster analysis for market definition purposes and the possibility of alternative outcomes also offers a way of testing rival positions on markets.

3.3   Hierarchical cluster analysis and its relation to market definition

The linkage method and distance metric are the similarity measures that dictate the type of grouping structure suggested by the cluster analysis. Having chosen these similarity measures, the analyst

can decide on the objective of the cluster analysis: the analyst may be interested either in strict classification where the number of groups are chosen *a priori* and cluster analysis is used to populate the groups or, alternatively, the analyst may be interested in an exploration of group structure without strong *a priori* views regarding the number of groups. Cluster analysts can therefore choose between hierarchical and non-hierarchical cluster analysis (see Kaufman and Rousseeuw (1990) and Gordon (1999) for comprehensive reviews). Non-hierarchical cluster analysis requires the analyst to specify an initial number of clusters and some type of starting configuration (Hastie, Tibshirani et al. 2001). This method generates the actual clusters as output. Hierarchical cluster analysis does not generate clusters, but produces a hierarchical tree diagram, called a dendrogram. The dendrogram consists of nodes, where each node represents a cluster and where each non-terminal node has two children nodes. At the lowest level, each observation represents a terminal node and the lower-level nodes are shown to successively merge at various "heights" until the entire data set is grouped as a single cluster at the top. The height of a merged node is proportional to the distance between the two children nodes, so that terminal nodes have zero height (Anderberg 1973; Everitt 1993).

Hierarchical clustering applies the following algorithm to generate the dendrogram (Johnson and Wichern 2002):

(i) Start with N nodes, each containing one data point.

(ii) Calculate a distance matrix in the dendrogram, $D = (d_{ik})$ (based on one of the distance concepts discussed earlier).

(iii) Search for the most similar pair of nodes in the distance matrix and merge these nodes to form a new dendrogram.

(iv) Calculate a new distance matrix for the new dendrogram, using the particular linkage method to calculate distances among original nodes and the newly formed node. The distance between node UV (i.e. containing element U and V) and another node W, $d_{(UV)W}$, is calculated as follows:

Average linkage: $d_{(UV)W} = \dfrac{\Sigma_i \Sigma_k d_{ik}}{N_{(UV)} N_W}$ where $N_{(UV)}$ is the number of elements in node UV and N_W is the number of elements in node W

Single linkage: $d_{(UV)W} = \min\{d_{UW}, d_{VW}\}$

Complete linkage: $d_{(UV)W} = \max\{d_{UW}, d_{VW}\}$

(v) Repeat step (iii) and (iv) until a single node remains.

Hierarchical clustering is preferable for market definition purposes for a number of reasons. Firstly, and perhaps most important, the analyst need not specify the number of clusters beforehand (Kaufman and Rousseeuw 1990). One purpose of the market definition enquiry is to decide on the number of markets, so that a technique that would force analysts to specify a number of markets *a priori* would be suboptimal.

Secondly – and related to the first benefit – hierarchical clustering is consistent with a broader view of the uses of market definition. Market definition involves a ranking of substitutes and this ranking is mirrored in the hierarchy of the dendrogram: by providing a hierarchical structure, substitution patterns can be assessed more effectively. The method clarifies whether products A, B and C, even if they are in a single cluster, are equally good substitutes. In fact, the dendrogram can be used to rank the substitutability of products even if no formal clusters are identified – thereby enabling one to identify the closest competitors of a particular product without forming clusters. The dendrogram, therefore, allows flexible boundary drawing: the analyst may start with the firm(s) under investigation and generate a set of concentric circles, where circles closer to the centre contain competitors closer to the firm(s). The relative ease with which one may interpret the dendrogram is a further benefit in this regard, as courts frequently find quantitative results hard to analyse.

Thirdly, the dendrogram prevents the competition analyst from forming small arbitrary markets without including other relevant competitors. The analyst may still elect to disregard the optimal number of clusters (defined below) in favour of more or fewer clusters, but in both cases the dendrogram forces the analyst to specify the height at which to form such clusters – in effect ensuring that the same criterion is applied to all clusters. Again, one market cannot be arbitrarily inflated

by a respondent or arbitrarily minimised by a complainant. In other words, the dendrogram ensures consistency in cluster formation. This property is particularly useful for market definition. In merger investigations, competition authorities are frequently less interested in the exact contours of the market and more in the question of whether two merging parties are in the same market. Put differently, authorities would like to know which other competitors, in addition to the merging firms, should be included at minimum. The dendrogram helps authorities answer this question, as is illustrated in the empirical application later.

Fourthly, the hierarchical cluster analysis described above is agglomerative: the analysis commences with a large number of nodes, merging successive nodes until a single node is formed. This agglomeration is consistent with the SSNIP test, which starts with a market containing a single firm and then adds competitors satisfying the SSNIP to this market. The dendrogram output is a graphical representation of this agglomeration, which enables practitioners to follow the SSNIP logic in a graphical fashion.

Finally, hierarchical clustering does not depend on a particular probability distribution and can be applied to cases where samples are rather small – a common occurrence in many competition investigations where the number of products to compare is quite few.

The benefits described above suggest that hierarchical cluster analysis is useful in exploring the hierarchical structure of substitution during the market definition exercise. However, apart from obtaining information about hierarchical substitution patterns, the competition analyst may also be concerned about drawing the actual market boundaries. Identifying the optimal number of markets or clusters from the dendrogram requires an optimality criterion, which allows the analyst to judge whether the clusters meet an *a priori* definition of cluster (Tibshirani, Walther et al. 2001). The following section explores optimality criteria and techniques for identifying the optimal number of markets.

3.4   Forming clusters from the dendogram

The process of identifying and forming clusters from the dendrogram is called tree cutting, as clusters are formed, in a graphical sense, by "cutting" the dendrogram at a particular height. More formally, tree cutting refers to halting the clustering algorithm at a particular height and taking the nodes at that height as the number of clusters. These clusters then represent the different markets. To avoid arbitrary delineation of markets, a criterion is required for choosing the optimal height for tree cutting.

The optimal height for tree cutting must be determined on a case-by-case basis: cluster analysis for different datasets generates different dendrograms and a fixed cut-off point in terms of height is not appropriate. An optimal clustering procedure should be true to the idea of clusters as densely packed groups of observations. Formally, an optimal clustering procedure should result in lower variance within clusters compared to the variance of uniformly distributed data grouped into the same number of clusters. This intuitive property is formalised in the gap statistic developed by Tibshirani, Walther et al. .(2001) to test a range of hypotheses regarding the number of clusters.

Following their notation, let D_r be the within-cluster variance, defined as the sum of all pair-wise distances for all observations within the rth cluster C_r:

$$D_r = \sum_{i_1, i_2 \epsilon C_r} d(y_{i_1}, y_{i_2})$$

(4)

Let k be the number of clusters and let W_k be a scaled version of the pooled within-cluster variance across all clusters:

$$W_k = \sum_{r=1}^{k} \frac{1}{2n_r} D_r$$

(5)

W_k is generally *de*creasing in k: as the estimated number of clusters increases and approaches the number of groups present in the data, within-cluster variance will decline by a significant margin.

Beyond the optimal number of clusters W_k continues to decline, but now at a much slower pace: splitting a single group in the data into two clusters reduces variance by much less than splitting two different groups in the data into two clusters. The slowdown in the W_k curve usually shows up as a "kink" at some value for k, say k^*, which then represents the optimal number of clusters (Hastie, Tibshirani et al. 2001).

The graphical approach to identify a kink can be formalised by the gap statistic developed by Tibshirani et al. (2001), which tests the null hypothesis of no clustering ($k = 0$) against the alternative of clustering ($k > 0$). For this purpose, let $E^*([\log(W)_k])$ be the expected pooled inter-cluster variance under the null, i.e. if the input data is assumed uniformly distributed. The optimal number of clusters is the point where the difference (or "gap") between W_k and $E^*(W_k)$ is at a maximum and is found by inspecting the gap statistic:

$$gap(k) = E^*([\log(W)]_k)) - \log(W_k) \tag{6}$$

The gap statistic outperforms alternative methods for finding the optimal number of clusters and performs well in cases where the data are not naturally separated into more than one group .(Tibshirani, Walther et al. 2001). Despite these advantages, the gap statistic is not monotonically decreasing where clusters contain well-separated sub-clusters. The non-monotone behaviour warns against a simplified search for a global maximum and suggests that it is appropriate to study the gap statistic graph to identify local maxima. Under these conditions different people will find different numbers of clusters optimal and the choice will be determined by the goal of the analysis, as demonstrated in the application.

This study implements cluster analysis using the R statistical language, which is an open-source object-oriented language and is frequently used for cluster analysis in the biological sciences (R Development Core Team 2009). The analysis relies heavily on the *hclust*, *pvclust* (Suzuki and Shimodaira 2009) and *clusterSim* (Walesiak and Dudek 2009) packages written for the R language. Standard statistical packages used by economists, including STATA, also support cluster analysis.

3.5 Robust dendrograms via bootstrapping

The dendrogram provides useful information on substitution patterns and can be used to identify an optimal number of markets. One remaining concern is whether any particular dendrogram is robust (Tukey 1977): if the competition practitioner identifies substitutes using a transparent loss function and grouping methodology as described above, the practitioner also seeks assurance that sampling error in the data has not compromised the dendrogram, in whole or in part. The search for robust, generalizable conclusions in competition policy is not new, as the growth of econometrics in empirical industrial organization shows (Baker and Rubinfeld 1999; Baker and Bresnahan 2008). The need for robust dendrograms is particularly acute in the biological sciences and it is this literature that has motivated the developments of techniques that has moved cluster analysis from a descriptive tool towards one that also allows statistical inference (Suzuki and Shimodaira 2006).

Felsenstein (1985) first suggested bootstrapping techniques to test whether a clustering pattern holds under reasonable assumptions. Bootstrapping is appropriate for cluster analysis, as cluster analysis does not require distributional assumptions concerning the input data: as the true population distribution is not necessarily known, it is not possible to derive exact test statistics. A data-based method, such as bootstrapping, overcomes the problem by treating the sample distribution as the population distribution. The procedure starts by selecting n_b observations from the data and performing a cluster analysis. The procedure is repeated for a large number of times and the bootstrap sample selection takes place with replacement (so that all of the original observations are available for selection for each new bootstrap sample). The bootstrap probability (BP) for a cluster or node in a dendrogram can then be calculated as the frequency with which the dendrogram or cluster appears in the bootstrap samples (Felsenstein 1985; Efron, Halloran et al. 1996). Subsequent work in this field has focused on the statistical size problem encountered in bootstrapping because

of the shape of the boundaries of different regions or clusters, showing a first-order size bias $O(n^{-1})$ (Shimodaira 2002).

The literature on bootstrapping for hierarchical clustering developed a range of alternative bootstrap techniques, including multi-scale bootstrapping, to reduce Type I error in the bootstrap probabilities. The probability values for which bias has been reduced are known as "approximately" unbiased (AU). AU estimates suffer from second-order size bias $O(n^{-\frac{1}{2}})$. The size improvement is particularly important in the current context, as the sample size is small: for $n = 20$, the size distortion for the BP is of an order $n^{-\frac{1}{2}} = 0.22$ and for the AU of an order $n^{-1} = 0.05$. The multi-scale bootstrapping procedure is AU under the condition that the dataset can be transformed to normality. In the case of normality, the curvature fitting (from which the size correction is derived) does not take into account that the covariance matrix may also vary. An improved multi-step version does not attempt to estimate the size correction directly as the two-level tries to, but instead generates bootstraps for different sample sizes (meaning sample sizes up to the actual size of the sample) and then for each sample size attempts to back out the implied curvature (by fitting a regression to the bootstrap probabilities). These calculations provide a third-order unbiased probability (Shimodaira 2004). In the current context, the multi-step multi-scale bootstrap, has a third-order size distortion of magnitude $n^{-\frac{3}{2}} = 0.01$.

The following section presents a South African merger case, which is used to demonstrate how hierarchical clustering analysis can assist in market definition.

4 Illustrative case: Primedia / KayaFM partial radio station merger

4.1 Validity of the case

In 2006 the South African media group Primedia together with an investment company Capricorn applied for approval from competition authorities to acquire a partial stake of 24.9% in the provincial radio station Kaya FM (Kaya hereafter). Primedia already owned two other radio stations in the same province, Highveld Stereo (Highveld) and Radio 702 (702), and the competition investigation revolved around market definition, as counterparties disagreed on the extent of the relevant product and geographic markets.

This case is particularly appropriate to illustrate how consumer characteristics data can assist in market delineation. Radio stations sell access to a market for information and entertainment goods. In this setup the listener is looking for access to this marketplace because she has specific preferences for news products, music products, talk shows, etc., and she pays with the opportunity cost of her time. Content is therefore uniquely aimed at a particular demographic or income group (Dewenter 2003; Lindstadt 2009): media consumers desire variety and may have specific preferences for language, culture, etc. The advertiser is looking for access to the marketplace so that she may sell an information product in the form of advertisements.

The radio station is therefore a market maker for information products, matching those seeking entertainment and information with those looking to provide those information goods and resources. Put differently, the radio station acts as a two-sided platform matching an audience of listeners with a group of advertisers (Argentesi and Filistrucchi 2007; Evans 2009). This feature of the market was accepted by all parties in the merger case. In a two-sided setting, it is important to note that consumer characteristics, as reflected in the demographic and income profile of listeners, carry significant information. The feature of the market important to advertisers is the listener profile, while the feature of the market important to listeners is information and entertainment, which, in turn, can *also* be approximated by the listener profile. In other words, consumer characteristics shed light on both sides of the market. This implies that using consumer characteristics to delineate markets in a two-sided media setting does not imply a sole focus on the demand behaviour of one side

(the listener side). Consumer characteristics become a broader measure of overall information market characteristics, providing information on both the listener and advertiser side. In effect, consumer characteristics become a type of proxy for product characteristics in this two-sided setting[2].

Of course, it is always useful to obtain direct evidence on both sides, as the approximation is imperfect (Filistrucchi 2008). The framework developed by Argentesi and Filistrucchi (2007) suggests that an assessment of market power in a two-sided radio market requires four pieces of information: (i) the price elasticity of demand of listeners (with respect to a notional price of listening), (ii) the price elasticity of demand of advertisers, (iii) the elasticity of demand of listeners with respect to quantity of advertising, and (iv) the elasticity of demand of advertisers with respect to quantity and quality of listeners. Based on our argument above, consumer characteristics shed light on (i) and (iv) and should ideally be complemented by further information on (ii) and (iii).

Some have argued that consumer characteristics are less important given that audiences (in particular, South African radio audiences) do not pay to listen. However, monetary payment is not a prerequisite for the definition of a relevant market. As noted above, the consumer still pays with the opportunity cost of her time. Nevertheless, it is true that the absence of monetary payment creates empirical challenges (Europe Economics 2002):

"[t]hese features make the economic understanding of markets inherently difficult, since they require an analysis of competitive constraints... without any of the price and volume data that are required by some quantitative techniques for market definition".

The zero monetary price of radio content therefore requires alternative measurement of the "notional" price in order to infer cross-price elasticities, which may be difficult .(Europe Economics 2002). However, as argued above, the distance between consumer characteristics of the different radio stations can assist. The Primedia/Kaya radio merger is therefore a useful case to demonstrate how consumer characteristics can assist in market delineation.

4.2   Background

Up to the mid-1990s, the South African radio industry was highly regulated, with the majority of stations run by the state-owned South African Broadcasting Corporation (SABC). In 1996 the broadcasting sector was partially liberalised and six major SABC stations were privatised (including two major Gauteng[3] radio stations Highveld Stereo and Radio Jacaranda). In early 1997 eight new commercial radio licences were granted for broadcasting in South Africa's three biggest cities (Johannesburg, Cape Town and Durban). The remaining SABC radio stations were also restructured and split into a commercial and a public group, with SABC commercial stations operating on the same profit basis as privately owned stations and SABC public stations operating on a non-profit basis and subsidised by the commercial stations.

Since the late 1990s, broadcasting authorities have also issued more than 100 licences for community radio stations run on a non-profit basis. However, the role of community-driven radio stations in constraining the exercise of market power did not receive significant attention in this merger investigation, because their listener numbers are generally small. Since these stations do not fundamentally alter the market definition conclusions obtained from the statistical analysis, the study does not consider them further.

Radio stations broadcasting in Gauteng can therefore be categorised into SABC commercial, SABC public and private commercial radio stations:

- SABC commercial: 5fm, Metro

---

[2]The author thanks Lapo Filistrucchi and other participants for raising this issue in a discussion of the paper at a conference.

[3]Gauteng is one of South Africa's nine provinces and the economic and political centre of the country, containing the cities of Johannesburg and Pretoria.

- SABC public: Ikwekwezi, Lesedi, Ligwalagwala, Lotus, Motsweding, Munghana Lonene, Phalaphala, Radio 2000, RSG, SAFM, Thobela, Ukhozi, Umhlobo Wenene

- Private commercial: 702, Highveld, Classic, Jacaranda, Kaya

The literature on radio market definition suggests that the type of content generally dictates whether the geographic market will be local or broader, partly because radio station licences may stipulate specific amounts of local or sub-national content (Trade Practices Commission 1994; Indepen 2004; Sandoval 2006). Geographic market definition by at least two of the parties in the case involved both regional and national markets: "Gauteng regional" markets for radio content customised to Gauteng audiences and "Gauteng national" markets for general national content available to Gauteng audiences[4]. If Gauteng radio listeners prefer local content, then radio stations with a national footprint and local stations without a Gauteng footprint must be excluded ("footprint" refers to the proportion of Gauteng province reached by the station signal). However, if general content is preferred, national radio stations with a Gauteng footprint are also plausible competitors for listeners. This study focuses on the broader geographic market that includes local radio stations as well as national radio stations with a Gauteng footprint. The complainant argued against the existence of a single Gauteng geographic market on the basis that radio station footprints do not overlap significantly (Competition Tribunal 2008: 10-11). However, the issue for market definition is sufficient and not significant overlap: as long as a sufficient number of listeners can switch to the particular station the SSNIP is not profitable.

As radio stations form part of the media industry, the first relevant question in the product market definition exercise would be whether to include radio as part of a larger media market. This issue has been particularly pertinent in recent years, given technological convergence (including the availability of radio stations online, potentially making them competitors for other media forms such as online newspapers and other websites). The literature is divided on this issue. On the one hand, there are strong indications of significant cross-media substitution. The evidence generally comes from the developed world, including US and UK studies (Silk, Klein et al. 2002; Waldfogel 2002). However, despite these claims, the dominant position in the competition literature is that radio markets are to be distinguished from other media markets, by virtue of their particularly personal and intrusive nature. In addition, authors working on the advertising side of radio markets also find radio advertising to constitute a separate market (Ekelund, Ford et al. 1999). In the current case, we will focus on radio as a distinct product market, following the approach of the competition authorities .(Competition Tribunal 2008: 10).

4.3   Data

The study relies on a dataset of radio listener characteristics obtained from the South African Advertising Research Foundation's Radio Audience Measurement Survey (RAMS) for 2004/5, the year in which the merger application was first heard before South African competition authorities (South African Advertising Research Foundation 2005). The RAMS is an extensive media research tool designed to help South African radio stations target their listener profiles more accurately. It is collected from the same respondents queried in the construction of the All Media and Products Survey (AMPS), which is a primary source of regularly collected data on South African consumer behaviour (Van Aardt 2008). The RAMS is constructed by means of "radio diaries" completed by all members of a particular target household, in which each member records, *inter alia*, the radio station(s) he/she has listened to during the seven days preceding the survey, including the corresponding times of the day and the duration.

The RAMS data is useful for this study, as it contains data on the demographic and income profile of listeners of each radio station. A survey respondent is defined as a listener of a particular radio

---

[4]The investigation did not consider Internet radio.

station if he/she has recorded listening to that station in the seven days preceding the survey. It is possible to consider separate listener profiles for different hours of the day, given that the audience for a breakfast show may differ from that for a late morning talk show. We do not attempt such disaggregated analysis and choose to focus on the overall number of listeners over seven days. This issue did not feature prominently in the merger proceedings.

The RAMS data contain the number of listeners of each radio station falling into each of the following categories (see Table 1).

The RAMS data do not measure or report listener income explicitly, but uses the LSM (Living Standards Measure) membership of a particular listener as an approximate indicator of the income group to which that listener belongs. The LSM structure is a 10-group measurement of the living standard of the South African population based on non-demographic variables – specifically, using ownership and access to a range of appliances and household amenities to infer income levels.

Figure 1 shows that the LSM group variable is a good indicator of income distribution. The subsequent analysis focuses on the LSM 6 to LSM 10 groups. The different parties to the merger agreed to focus on this cohort, although, as is discussed below, they differed on whether to consider this group as a whole for market definition purposes.

The RAMS data capture these demographic and income characteristics in *absolute* numbers (thousands of listeners). However, consumer characteristics are better measured in proportions. Absolute numbers will bias market definition using cluster analysis, as radio stations with large audiences will be found not substitutable with radio stations with small audiences, despite the possibility that the smaller radio station is catering for a similar audience. Appendix A reports an extract from the actual RAMS data and shows how these numbers are converted to proportions.

It is important to ensure that the consumer characteristics that are observed are representative of the true consumer characteristics that matter for the SSNIP – this ensures that the empirical analysis is relevant for market definition purposes. Suppose one considers race and age as the only dimensions on which to compare consumers and, therefore, uncovers significant differences between old white and young black listeners; if income and sex, rather than race and age, are the more relevant drivers of demand, such a comparison is misleading. Berry, Levinsohn and Pakes (2004) argue that it is difficult to identify the salient attributes of products, as the analyst has no way of knowing that all the salient features have been identified. This difficulty frequently leads to contention in competition cases surrounding the choice of particular variables.

In the radio case discussed here, parties did not disagree about the inclusion of specific variables. Nevertheless, three comments are appropriate. Firstly, the variables were obtained from an advertising industry body which can be assumed to be collecting data relevant to its members for targeting purposes and which confirms the salience of these variables in a demand analysis. Secondly, extant literature provides guidance. In the spirit of the general-to-specific modelling approach, it is necessary for empirical models to be congruent with theory, that is, to account for all previous research findings. The broader competition policy literature related to radio station mergers and investigations include a range of cultural and ethnic variables – note for example the recent US discussions on Hispanic radio station content (Sandoval 2006; Coffey and Sanders 2010). Thirdly, data availability also constrains the set of available variables.

In the current case, the disagreement between the applicant and defendant related less to the actual inclusion of specific variables (such as race or age) and more to the level of aggregation of the individual income variables. Whereas the applicant preferred a broad income variable that includes all LSM 6-10 listeners, the defendant argued in favour of a more nuanced approach. The decision on which level of aggregation is appropriate should be driven by the aim, which is to assess substitution patterns. As Berry et al. (1995: 852) note, the major restriction of the earlier models of differentiation (including Bresnahan (1981)) is the limited set of characteristics, which restricts the substitution patterns. It seems lop-sided to impose too many constraints on the number of characteristics, where the focus is to measure similarity. Although a finer income distribution will result in finer distinctions, it is always possible to evaluate the resulting clusters at a higher

level of aggregation. However, the benefit of starting with more rather than fewer variables lies in multiple variables enabling one to form a hierarchical picture of substitution, i.e. to assess how close competitors are rather than simply identifying a number of potential competitors.

4.4 Graphical assessment of data

Figure 2 shows the distribution of the consumer characteristics variables for the three merging stations. The figure is a first step in an attempt to group radio stations into markets on the basis of their consumer characteristics and highlights the income and racial differences, especially between Highveld and Kaya listeners. However, although the stations differ significantly on income and race, they are more similar on the basis of age – especially when compared to 702. This illustrates the problem of integrating diverging univariate conclusions into a single conclusion on multivariate similarity. Furthermore, it is not clear what is meant by similarity in this graphical assessment. Grouping based on graphical output is arbitrary at best and it is for this reason that the study suggests cluster analysis, a systematic grouping method that can also deal with multivariate observations.

5 Results from cluster analysis

This section shows the results of applying hierarchical cluster analysis to the radio listener characteristics data described in the previous section. Results are reported separately for the average, single and complete linkage methods.

5.1 Average linkage method

Figure 3 presents the dendrogram for radio listener characteristics based on the average linkage method. In addition to the different radio stations, the dendrogram reports the bootstrap probabilities (approximately unbiased).

The dendrogram provides important evidence for market definition purposes. The shape of the dendrogram suggests that there are radio stations difficult to group, shown on the left-hand side of the dendrogram. These are stations with specific types of audiences: for example, 702Talk (talk radio station), Lotus (a station with predominantly Indian listeners), Classic (a classical music radio station) and four other stations. However, more important, the dendrogram also identifies a group of radio stations that are quite similar in listener profile, clustering in a single node on the right-hand side of the dendrogram – starting with Kaya and ending with RSG. This node has important market definition implications. As noted, Ikwekwezi and Motsweding form a cluster first, followed by Jacaranda & RSG, then Highveld and Lesedi, and then Kaya and Metro. Importantly, Kaya and Highveld only cluster together at a late stage (stage 12), suggesting that a number of competitors separate these two radio stations.

One of the salient market definition issues in the current case was precisely which stations to include, *at minimum*, with Highveld and Kaya, if one assumes that Highveld and Kaya share the same market. One can ask the same question for 702 (instead of Highveld). However, this chapter focuses on the relationship between Highveld and Kaya rather than 702 and Kaya, as the former received most of the attention in the competition proceedings. This follows from Highveld's reputation as a "must have" station. Figure 3 suggests that, if Highveld and Kaya share a market, a number of other stations *have* to be considered: Metro, 5FM, YFM, Munghana Lonene, Lesedi, Ukhozi, Umhlobo Wenene, Ikwekwezi, Motsweding, Jacaranda and RSG.

As argued earlier, bootstrap techniques allow the analyst to assess the robustness of the dendrogram with respect to small data errors. The AU bootstrap probabilities in Figure 3 suggest generally low levels of confidence in large parts of the dendrogram. This indicates the difficulty of identifying all radio stations that should be included if Kaya and Highveld are in the same market,

as the same set of competitors are not likely to be identified if the data underwent small changes. Instead, it is better to identify those parts of the dendrogram that are more robust. In the current case, a cluster containing both Kaya and Highveld is not necessarily robust. However, it is clear that Highveld shares a robust cluster with a set of other competitors – indicating that it is important first to include these other competitors in Highveld's market before arguing for the inclusion of Kaya.

The dendrogram not only provides information about substitution patterns, but also assists in drawing exact market boundaries. As argued earlier, the gap statistic can be used to identify the optimal number of markets. The gap statistic in this study requires random draws from the uniform distribution, which is generated using a Monte Carlo simulation of 1000 repetitions. Figure 5 plots the gap statistic for each assumed number of clusters.

The gap statistic suggests that the data is not easily separable into clearly identifiable clusters: the gap statistic continues to rise as the number of clusters increases. Nevertheless, the gap statistic does indicate a maximum at two clusters, as it does not increase significantly when adding a third, fourth or fifth clusters. Given the hierarchical structure, this plateau effectively implies that all but one station are best grouped into the same broad market. However, two clusters represent a local maximum, as the plateau from two to five clusters is followed by a sharp and persistent increase in the gap statistic. Such a shape indicates that there are a number of well-defined sub-clusters within the larger clusters (Tibshirani, Walther et al. 2001). Therefore, if the analyst is interested in identifying a large number of tight clusters, he should consider the global maximum, which suggests 18 or more clusters: the stations should therefore be grouped into a large number of markets, each containing only one radio station. Nevertheless, competition analysts are likely to be more interested in *local* maxima, as local maxima shed light on broader grouping patterns. This requires inspection of the entire graph to identify local maxima rather than merely searching the graph for the global maximum.

The dendrogram presented above is based on the average linkage method and it is useful to compare the outcomes for the single and complete linkage methods, as shown in the following subsections.

5.2   Single linkage method

Figure 6 repeats the dendrogram using the single linkage method.

The results for the single and average linkage methods are similar. Figure 6 suggests the same substitutes identified by Figure 3 for a market in which both Highveld and Kaya compete: Ukhozi, Munghana Lonene, 5FM, Lesedi, Highveld, Umhlobo Wenene, RSG, Jacaranda, Ikwekwezi, Motsweding, YFM, Kaya and Metro. The boundaries of the relevant market are therefore defined identically for the single and average linkage methods. However, bootstrap results suggest that the single linkage method dendrogram is more robust than the dendrogram obtained for the average linkage method; the larger clusters in Figure 6, including the one containing both Highveld and Kaya, have much higher bootstrap probabilities for the single method.

To identify the optimal number of clusters, the gap statistic is reported in Figure 7:

Similar to the average linkage method, the gap statistic also seems to grow quite slowly for clusters 2 through 5 – but do not reach a local maximum. Figure 7 suggests a local maximum at thirteen clusters: the gap statistic rises significantly from 2 through to 12, but declines for 13. Moving beyond 13 the gap statistic rises significantly again to reach a global maximum at around 18 clusters. These outcomes suggest that a mechanical interpretation of the gap statistic should be avoided: if the plot gives conflicting signals it is likely that the data, under this linkage method, does not form well-separated groups. Where it is difficult to separate data into distinct groups, it is advisable to rely on the dendrogram to answer questions about substitution patterns rather than attempt to group radio stations into a specific number of clusters.

5.3   Complete linkage method

Figure 8 presents the dendrogram for the complete linkage method.

This dendrogram is quite different from both the average or single linkage methods. The dendrogram suggests two large clusters, one containing both Highveld and 702 and the other containing Kaya. The bootstrap probabilities in this case suggest low statistical confidence in the clusters, even if they are formed using an optimality criterion. In this case, similar to the average linkage method, it is preferable to consider sub-clusters that do have high bootstrap confidence levels. As far as Highveld is concerned, Figure 9 suggests that Jacaranda, RSG, Highveld and Lesedi should be grouped (79% probability), while Kaya and Metro appear close (67%). This suggests that Kaya and Highveld cannot be considered without considering these stations – although these do not represent the minimum number. These are the stations with the best indications of substitutability.

To identify the optimal number of clusters, the gap statistic is reported in Figure 9:

The specific form of this dendrogram suggests that all radio stations have to be considered if Kaya and Highveld are grouped together. The gap statistic appears to be support this and does not reach any local maximum, although the statistic grows slowly from two to three clusters.

5.4   The gap statistic and uncertainty

This study calculates gap statistics using Monte Carlo simulations and the analyst should account for the simulation error when determining the optimal number of clusters. Tibshirani et al. (2001) suggest the '1-standard deviation rule', where the number of clusters, k, is chosen as the smallest k such that

$$gap(k) \geqslant gap(k+1) \tag{7}$$

where $s(k) = \frac{sd(k)}{\sqrt{1+\frac{1}{B}}}$ and $sd(k)$ is the standard deviation of W_k and B is the number of Monte Carlo repetitions. The '1-standard deviation rule' follows Breiman, Friedman, Olshen and Stone (1984), who consider rules of thumb for selecting the optimal . Figures A1, A2 and A3 in Appendix A plot the value of $gap(k) - [gap(k+1) - s(k+1)]$ against k for each of the linkage methods. The calculations suggest two as the optimal number of clusters, which effectively implies that the data is quite difficult to separate into distinct clusters. This underlines the earlier point that the data for this case is better investigated asking specific substitution questions rather than attempting to delineate specific markets.

6   Summary and discussion of cluster results

6.1   Summary of results

The dendrograms presented above are useful graphical tools that capture a significant amount of information on substitutability. This section summarises the conclusions from each linkage method in a single table. In particular, Table 2 shows which other stations would populate a market that included both Highveld and Kaya. The table also allows one to assess how robust the suggested markets are, if one is to limit conclusions to the *largest* possible clusters with an 80% or higher bootstrap confidence level.

Table 2 shows that the average and single linkage methods suggest the same substitutes. However, the highlighted areas suggest (as noted earlier) that the overall conclusions are stronger for the single linkage method – there is a more than 80% probability (in a bootstrap sense) that the entire cluster will be reproduced exactly in the event of small random changes to the data. Nevertheless, the results for the average linkage method are also strong: although there is uncertainty about the larger cluster in its entirety, two large sub-clusters are likely to be reproduced (labelled "(1)" and "(2)"). The position of Kaya is less certain: it may be that under small data changes Kaya does

not appear in this cluster. However, even if one insists that Kaya should be included, the results suggest that before including Kaya with Highveld, the set of stations in the (1) group will have to be considered first.

The results for the complete linkage method are quite different. Firstly, the dendrogram suggests a much larger number of competitors. Secondly, there is considerable uncertainty, which makes this particular cluster less useful, although there are certain sub-clusters that mirror the sub-clusters identified from the average linkage method (Highveld and Lesedi are again grouped together, and the group (2) stations under the average method also appears to organize into robust clusters under the complete method).

It is important to note that this suggested interpretation of the relevant market mirrors the decision of the Competition Tribunal, who identified an "inner circle" of closer competitors and an "outer circle" of less-close competitors.

The cluster analysis produces the same conclusions on market definition using the average and single linkage method (although statistical confidence is higher for the single linkage), but suggests a larger market using the complete linkage method. Reconciling these pieces of evidence is an important challenge to market definition, but not a unique challenge to cluster analysis. Econometric models can produce quite different price elasticity estimates and hence potential relevant markets, or as is shown in the following chapter, price tests may suggest a variety of potential relevant markets. The outcomes under each depend on the particular question. Also, it depends on the extent to which the systematic quantitative evidence matches other pieces of evidence. The chapter on market definition as a problem of statistical inference deals explicitly with the issue of uncertainty and diversity of evidence in market definition.

6.2   Corroborating evidence

The cluster analysis of radio listener characteristics offers an alternative tool for the delineation of radio listener markets. As this specific tool is chosen precisely because of data constraints, it is generally difficult to find alternative evidence (at least quantitative evidence) from the listener side. As mentioned earlier, radio markets are two-sided and the analyst must also study the advertiser side of radio market. Both sides were studied during the merger hearings, but in this study the interest lies with the extent of the listener side of the market in particular. However, other data from RAMS offer potentially useful corroborating evidence. The RAMS data suggest that, on average, South African radio listeners tune in to two to three radio stations per day. In fact, the RAMS data enable the analyst to calculate which proportion of listeners of a particular radio station listen to each of the other radio stations. These switching proportions then offer an alternative ranking of cross-price elasticity between the different radio stations, as shown in Table 3.

Table 3 confirms the findings from the cluster analysis. Firstly, for Highveld, Kaya and 702, the substitutes identified by the cluster analysis and the substitutes from the switching analysis are quite close. Secondly, it also confirms the conclusion that, if one includes Highveld and Kaya in the same market one would have to consider a significantly wider range of potential substitutes with these two stations.

The cluster analysis suggests that Highveld and Kaya may be in the same overall market, but the dendrogram show that they are not necessarily close competitors and that the analyst should consider a range of competitors for Highveld other than Kaya. While these conclusions stem from a demand-side analysis of consumer characteristics, they appear consistent with supply-side evidence. The South African radio industry is a highly regulated industry and radio station licences are granted and administered by the Independent Communications Authority of SA (ICASA). Both the type of music and the language to be used are usually specified in the licence agreement. Although both Highveld and Kaya hold licences to play "adult contemporary" music, the Kaya licence constrains the format to be focused on an African audience with a particular mixture of music and talk. The regulatory environment therefore creates a differentiated market which restricts the extent to which

certain radio stations can compete for customers.

7 Limits of cluster analysis and consumer characteristics as market definition tools

A number of criticisms may be raised against the use of cluster analysis and consumer characteristics for market definition. Before considering general criticisms of cluster analysis, it is useful to note specifically its limitations in delineating media markets. In a recent study, Argentesi and Filistrucchi .(2010) argue that some media markets are a special type of two-sided market where indirect network externalities[5] only go in one direction: they find, for example, that the amount of newspaper advertising does not have a significant effect on the number of newspaper readers. This feature implies that an SSNIP on the advertiser side need not have an indirect impact on readers – which may imply that the SSNIP test could yield different market sizes depending on the side on which it is applied. In this study, an analysis of consumer characteristics is assumed to yield a market definition that applies equally well to both listener and advertiser sides. Although, as argued before, this assumption has merit, the cluster analysis should be complemented by other pieces of evidence.

Beyond two-sided media markets, cluster analysis faces a number of general limitations as a tool for market definition. One obvious criticism of a comparison of consumer characteristics for market definition would be that it does not include any price information. However, the analysis of consumer heterogeneity represents one part of a broader investigation based on other qualitative evidence. Furthermore, where price evidence is quite limited, the analysis of consumer characteristics may be useful for market definition. In media markets where products are highly differentiated and where the demand function may differ even between different time slots and for different content, it is likely to be very difficult to obtain representative price series (quantity series in the form of listener numbers or profiles are less difficult to come by; see the recent study by Bjornerstedt and Verboven (2009) on the difficulty of obtaining representative prices for differentiated goods). Another – and related – defence would be that, in certain markets, price data is *never* available. In most media markets in particular, monetary transactions only take place on one side of the market: the radio listener does not pay the radio station, nor do viewers or readers of free television or newspapers. Under these market conditions it is much more appropriate to rely on proximate variables (including consumer characteristics) and develop a set of reasoned assumptions to infer substitutability rather than to discard such evidence. As shown in this study, the analysis of similarity of consumer profiles has a well-defined economic basis and can be conducted in a systematic fashion.

Some have made a more general argument that only sophisticated econometric techniques can assist in market definition and have argued against the use of, *inter alia*, price correlation tests and other simple econometric techniques (Scheffman and Spiller 1987; Massey 2000; Hosken and Taylor 2004). As argued earlier, the arguments in favour of using consumer characteristics in the delineation of markets are not intended to undermine the use of what are now conventional econometric techniques. This study presents a complement rather than a substitute. Furthermore, one should not discard some quantitative techniques because of their supposed simplicity. The question is relevance rather than simplicity. In the way presented here, the assumptions underlying an analysis of consumer characteristics is made clear and can be subjected to cross-examination and refuting evidence by the different parties to a merger. Besides, competition authorities do not necessarily assign one quantitative piece of evidence higher weight than others.

The argument against a single encompassing tool for market definition and in favour of a rich toolkit also responds to the other, and strongly related, criticism of cluster analysis as being a purely "statistical" tool, arguably referring to the fact that cluster analysis is a data-driven technique

---

[5]Two-sided markets are (partly) defined by the presence of "indirect network externalities". For example, there is a positive indirect network externality running from listeners to advertisers: the more listeners a radio station attracts, the more attractive it becomes to advertisers.

and does not incorporate economic theory *a priori*. The trade-off between data and theory in econometrics is an old issue and econometric tools across the spectrum, from purely data-driven toward strongly theory-based approaches, continue to be used (see Pagan (2003) for an exposition in a macro-econometric context). It is therefore important for analysts to be aware of the data-driven nature of cluster analysis and to investigate further the reasons for any similarity of consumer characteristics uncovered using cluster analysis.

The hypothetical monopolist test deals with the sensitivity of *marginal* rather than *average* consumers: the question is not whether most consumers will switch but whether a number of consumers will switch such that profitability will be reduced. Cluster analysis uses average consumer profile similarity of two competing products as a measure of their cross-price elasticity. The problem with an average profile is therefore that it focuses attention on the profile of the dominant or average consumer rather than the marginal consumer. This author has argued that the cluster analysis still allows one to assess to where marginal consumers may turn, as it is based on identifying substitutes for the product under investigation, on the basis of the similarity of products on some rather than all consumer characteristics. Nevertheless, the average profile may still result in the underestimation of the extent of substitution between some radio stations if the consumer most sensitive to price looks very different from the average consumer.

Finally, cluster analysis does not recognise that markets in competition policy are not necessarily unique or mutually exclusive. For example, Highveld may be a competitor in Kaya's market, but Kaya may not be a competitor in Highveld's market. The cluster analysis treats competition as inherently symmetric, which follows from the use of a symmetric distance concept.

8 Conclusion

Market definition is a tool to assist practitioners in assessing the competitive implication of a merger or a particular type of conduct. The operationalisation of the HM test remains a complex task that cannot be reduced to a single tool, however econometrically sophisticated. This is the result of the inherent limits of our econometric and economic models. Cluster analysis represents another addition to the toolkit of practitioners that wish to evaluate the similarity in characteristics of products. It is useful both in the conventional sense, i.e. for delineating markets for market share calculations, and under an effects-based competition regime, for the study of substitution patterns.

Cluster analysis is particularly useful as it, firstly, imposes consistency – forcing parties to include all relevant products – and, secondly, improves understanding of quantitative evidence by providing graphical evidence to the competition authorities and courts. The cluster analysis literature has also made significant advances in moving the technique from a purely exploratory tool towards enabling statistical inference, which further enhances its attractiveness as a systematic tool for competition policy.

References

[1] Anderberg, M. R. (1973). Cluster Analysis for Applications. New York, Academic Press.

[2] Argentesi, E. and L. Filistrucchi (2007). "Estimating market power in a two-sided market: the case of newspapers." Journal of Applied Econometrics **22**: 1247-1266.

[3] Argentesi, E. and L. Filistrucchi (2010). Market definition in a two-sided market: the case of newspapers. Competition and Regulation European Summer School and Conference 2010. Chania, Crete, Greece.

[4] Baker, J.B. and T.F. Bresnahan (2008). Economic Evidence in Antitrust: Defining Markets and Measuring Market Power. <u>Handbook of Antitrust Economics</u>. P. Buccirossi. Cambridge MA, MIT Press.

[5] Baker, J. B. and D. L. Rubinfeld (1999). "Empirical methods in antitrust litigation: review and critique." <u>Am Law Econ Rev</u> **1**(1): 386-435.

[6] Berry, S.J. Levinsohn, et al. (1995). "Automobile prices in market equilibrium." <u>Econometrica</u> **63**(July 1995): 841-890.

[7] Berry, S.J. Levinsohn, et al. (2004). "Estimating Differentiated Product Demand Systems from a Combination of Micro and Macro Data: The New Car Model." <u>Journal of Political Economy</u> **112**(1): 68-105.

[8] Berry, S.T. (1994). "Estimating discrete-choice models of product differentiation." <u>Rand Journal of Economics</u> **25**(2): 242-262.

[9] Bishop, S. and M. Walker (2002). The economics of EC competition law: concepts, application and measurement. London, Sweet and Maxwell. Bjornerstedt, J. and F. Verboven (2009). Predicting the price effects of a merger in the Swedish market for analgesics - a simulation analysis. Fourth European Conference on Competition and Regulation. Chania, Crete, Greece, CRESSE.

[10] Breiman, L., J. Friedman, et al. (1984). <u>Classification and regression trees</u>. Belmont, Wadsworth.

[11] Bresnahan, T. (1981). "Departures from Marginal-Cost Pricing in the American Automobile Industry: Estimates for 1977-1978." <u>Journal of Econometrics</u> **17**(2): 201-227.

[12] Coate, M. B. and J. H. Fischer (2008). "A practical guide to the hypothetical monopolist test for market definition." <u>Journal of Competition Law and Economics</u> **4**(4): 1031-1063.

[13] Coffey, A. J. and A. K. Sanders (2010). "Defining A Product Market for Spanish-Language Broadcast Media: Lessons from United States v. Univision Communications, Inc. and Hispanic Broadcasting." <u>Communication Law and Policy</u> **15**(1): 55 - 89.

[14] Competition Tribunal (2008). Primedia Ltd, Capricorn Capital Partners Pty Ltd, New Africa Investments Ltd v Competition Commission and African Media Entertainment Ltd. <u>39/AM/MAY06</u>, South African Competition Tribunal.

[15] Davis, P. and E. Garcés (2010). <u>Quantitative techniques for competition and antitrust analysis.</u> Princeton, Princeton University Press.

[16] Dewenter, R. (2003). The economics of media markets. <u>Helmut Schmidt University, Faculty Economics, Discussion Paper 10</u>

[17] Efron, B., E. Halloran, et al. (1996). "Bootstrap confidence levels for phylogenetic trees." <u>Proceedings of the National Academy of Sciences USA</u> **93**: 13, 429–13, 434.

[18] Ekelund, R. B., G. S. Ford, et al. (1999). "Is Radio Advertising a Distinct Local Market? An Empirical Analysis." <u>Review of Industrial Organization</u> **14**: 239-256.

[19] Europe Economics (2002). Market Definition in the Media Sector: Economic Issues. Brussels, European Commission: Director-General Competition.

[20] Evans, D. S. (2009). "Two-Sided Market Definition." ABA Section of Antitrust Law, Market definition in antitrust: theory and case studies. **Forthcoming**

[21] Everitt, B. S. (1993). Cluster Analysis. London, Edward Arnold.

[22] Farrell, J. and C. Shapiro (2010). "Recapture, Pass-Through, and Market Definition." Antitrust Law Journal **76**(3): 585-60.

[23] Felsenstein, J. (1985). "Confidence limits on phylogenies: an approach using the bootstrap." Evolution **39**: 783–791.

[24] Filistrucchi, L. (2008). "A SSNIP Test for Two-Sided Markets: The Case of Media." SSRN eLibrary

[25] Froeb, L. M. and G. J. Werden (1991). "Residual demand estimation for market delineation: complications and limitations." Review of Industrial Organization **6**: 33-48.

[26] Gordon, A. D. (1999). Classification. London, Chapman & Hall / CRC Press.

[27] Gorman, T. (1980). "A possible procedure for analysing quality differentials in the egg market." Review of Economic Studies **47**(5): 843-856.

[28] Hastie, T., R. Tibshirani, et al. (2001). The elements of statistical learning: data mining , inference and prediction. New York, Springer-Verlag.

[29] Hayek, F. A. (1946 [1984]). The Meaning of Competition Individualism and Economic Order. F. A. Hayek. Chicago, Chicago University Press.

[30] Hosken, D. and C. T. Taylor (2004). "Discussion of 'Using stationarity tests in antitrust market definition'." American Law and Economics Review **6**(2): 465-475.

[31] Indepen (2004). Radio Mergers Are Not Substantial: The impact of the Communications Act and the Enterprise Act on radio merger. London, Commercial Radio Companies Association.

[32] Ivaldi, M. and F. Verboven (2005). "Quantifying the effects from horizontal mergers in European competition policy." International Journal of Industrial Organization **23**: 669-691.

[33] Johnson, R. A. and D. W. Wichern (2002). Applied Multivariate Statistical Analysis. Upper Saddle River, Prentice-Hall.

[34] Johnson, S. C. (1967). "Hierarchical clustering schemes." Psychometrika **32**: 241-254.

[35] Katz, M. L. and C. Shapiro (2003). "Critical loss. Let's tell the whole story." Antitrust Magazine **Spring 2003**: 49-56.

[36] Kaufman, L. and P. J. Rousseeuw (1990). Finding groups in data: an introduction to cluster analysis. New York, Wiley.

[37] Kelly, C. and J. Rice (1990). "Monotone smoothing with application to dose-response curves and the assessment of synergism." Biometrics **46**: 1071-1085.

[38] Lancaster, K. (1966). "A new approach to consumer theory." The Journal of Political Economy **74**(2): 132-157.

[39] Lancaster, K. (1971). Consumer demand New York, Columbia University Press.

[40] Lance, G. N. and W. T. Williams (1967). "A general theory of classificatory sorting strategies: 1 Hierarchical Systems." The Computer Journal **9**(4): 373-380.

[41] Lindstadt, N. (2009). Multisided media markets: applying the theory of multisided markets to media markets. University of Southern Denmark, Department of Environmental and Business Economics, Working Paper No 96. Esbjerg.

[42] Lorr, M. (1983). Cluster Analysis for Social Scientists. London, Jossey-Bass.

[43] Manne, G. A. and J. D. Wright (2009). Innovation and the Limits of Antitrust. George Mason Law & Economics Research Paper No. 09-54, Available at SSRN: http://ssrn.com/abstract=1490849

[44] Massey, P. (2000). "Market definition and market power in competition analysis: some practical issues." The Economic and Social Review 31(4): 309-328.

[45] McFadden, D. (1981). Econometric Models of Probabilistic Choice. Structural Analysis of Discrete Data with Econometric Applications. C. F. Manski and D. McFadden. Cambridge, Mass, MIT Press.

[46] McNulty, P. J. (1967). "A note on the history of perfect competition." Journal of Political Economy 75(4): 395-399.

[47] Morgan, B. J. T. and A. P. G. Ray (1995). "Non-uniqueness and inversions in cluster analysis." Applied Statistics 44(1): 117-134.

[48] Motta, M. (2006). Competition policy. London, Sage.

[49] Nevo, A. (2001). "Measuring market power in the ready-to-eat cereal industry." Econometrica 69: 307-342.

[50] O'Brien, D. P. and A. L. Wickelgren (2003). "A critical analysis of critical loss analysis." Antitrust Law Journal 71: 161-184.

[51] Pagan, A. R. (2003). An examination of some tools for macro-econometric model building. Eighth African Econometrics Conference. Stellenbosch.

[52] Petrin, A. (2002). "Quantifying the benefits of new products: the case of the minivan." Journal of Political Economy 110(4): 705-729.

[53] R Development Core Team (2009). R: A language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

[54] Sandoval, C. J. K. (2006). "Antitrust Law on the Borderland of Language and Market Definition: Is There a Separate Spanish-Language Radio Market? A Case Study of the Merger of Univision and Hispanic Broadcasting Corporation." University of San Francisco Law Review 40: 1-68.

[55] Scheffman, D. T. and P. T. Spiller (1987). "Geographic market definition under the U.S. Department of Justice Meger Guidelines." Journal of Law and Economics 30: 123-147.

[56] Shimodaira, H. (2002). "An approximately unbiased test of phylogenetic tree selection." Systematic Biology 51: 492–508.

[57] Shimodaira, H. (2004). "Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling." The Annals of Statistics 32(6): 2616-2641.

[58] Silk, A. J., L. R. Klein, et al. (2002). "Intermedia substitutability and market demand by national advertisers." Review of Industrial Organization 20(4): 323-348.

[59] Sokal, R. R. and C. D. Michener (1958). "A statistical method for evaluating systematic relationships." University of Kansas Scientific Bulletin 38: 1409-1438.

[60] South African Advertising Research Foundation (2005). Radio Audience Measurement Survey.

[61] Suzuki, R. and H. Shimodaira (2006). "Pvclust: an R package for assessing the uncertainty in hierarchical clustering." Bioinformatics **22**(12): 1540-1542.

[62] Suzuki, R. and H. Shimodaira (2009). pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling.

[63] Tibshirani, R., G. Walther, et al. (2001). "Estimating the number of clusters in a data set via the gap statistic." Journal of the Royal Statistical Society, Series B, Statistical methodology **63**(2): 411-423.

[64] Trade Practices Commission (1994). Market definition and competition issues in commercial broadcast radio. Canberra, Australian Government Publishing Service.

[65] Tukey, J. W. (1977). Exploratory data analysis. Reading, MA, Addison-Wesley Publishing Company.

[66] United States Department of Justice and Federal Trade Commission (1992). Horizontal merger guidelines.

[67] United States Department of Justice and Federal Trade Commission (2010). Horizontal merger guidelines.

[68] Van Aardt, C. J. (2008). "Determining the predictors of living standards in South Africa: a real world econometric approach." Southern African Business Review **12**(1): 1-17.

[69] Waldfogel, J. (2002). Consumer substitution among media. Media Bureau Staff Research Papers 2002-2003

[70] Walesiak, M. and A. Dudek (2009). clusterSim: Searching for optimal clustering procedure for a data set.

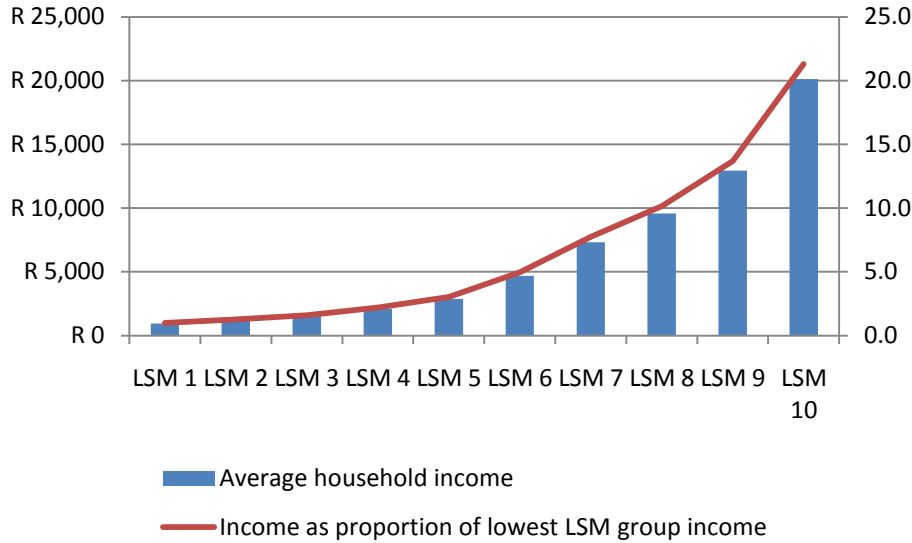Figure 1. LSM group as income group proxy (income in South African rand (R)).

Figure 2. Listener profiles per LSM (income), age, race and sex for merging radio stations
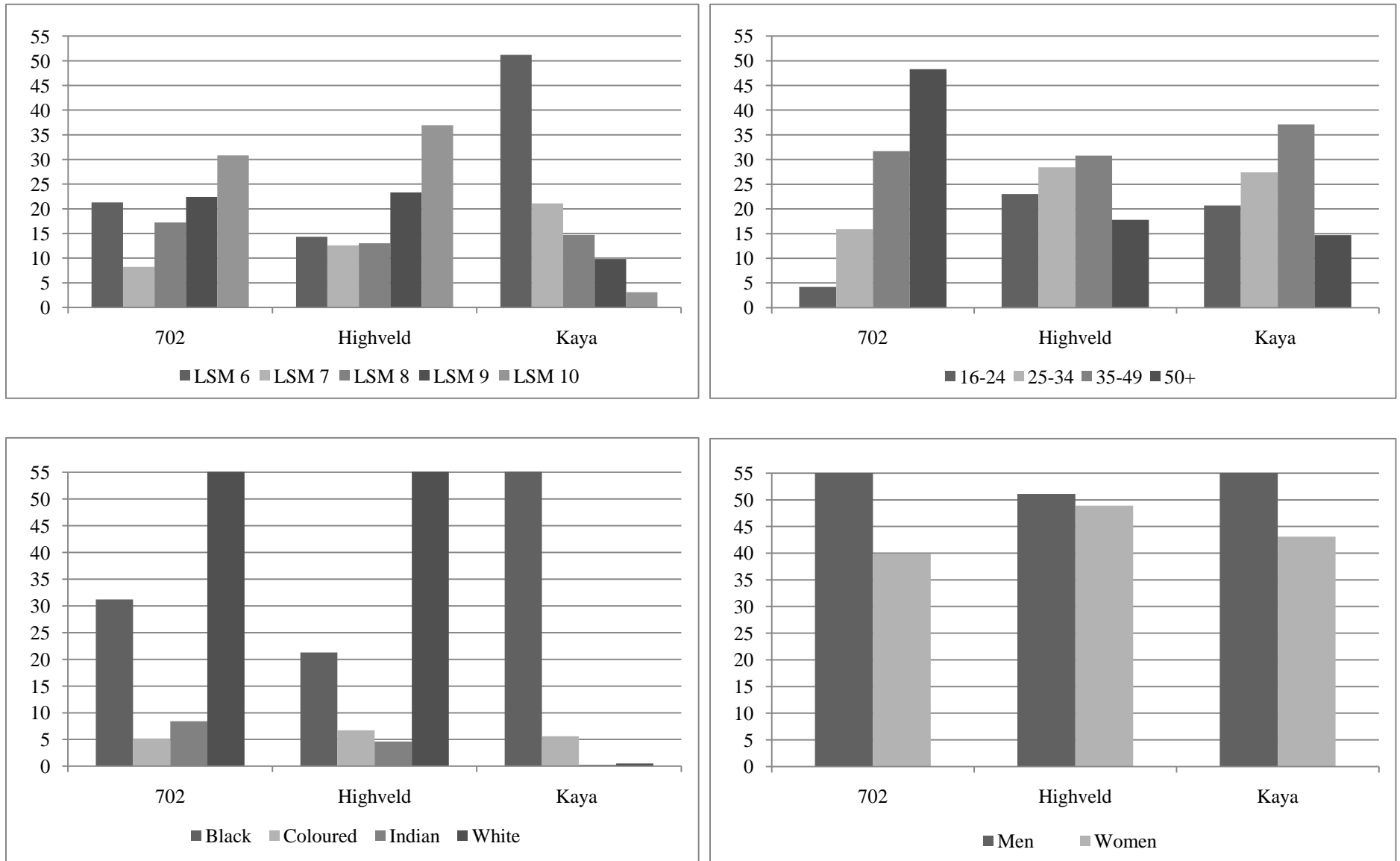
Figure 3. Dendrogram based on radio listener characteristics, average linkage method and statistical distance.
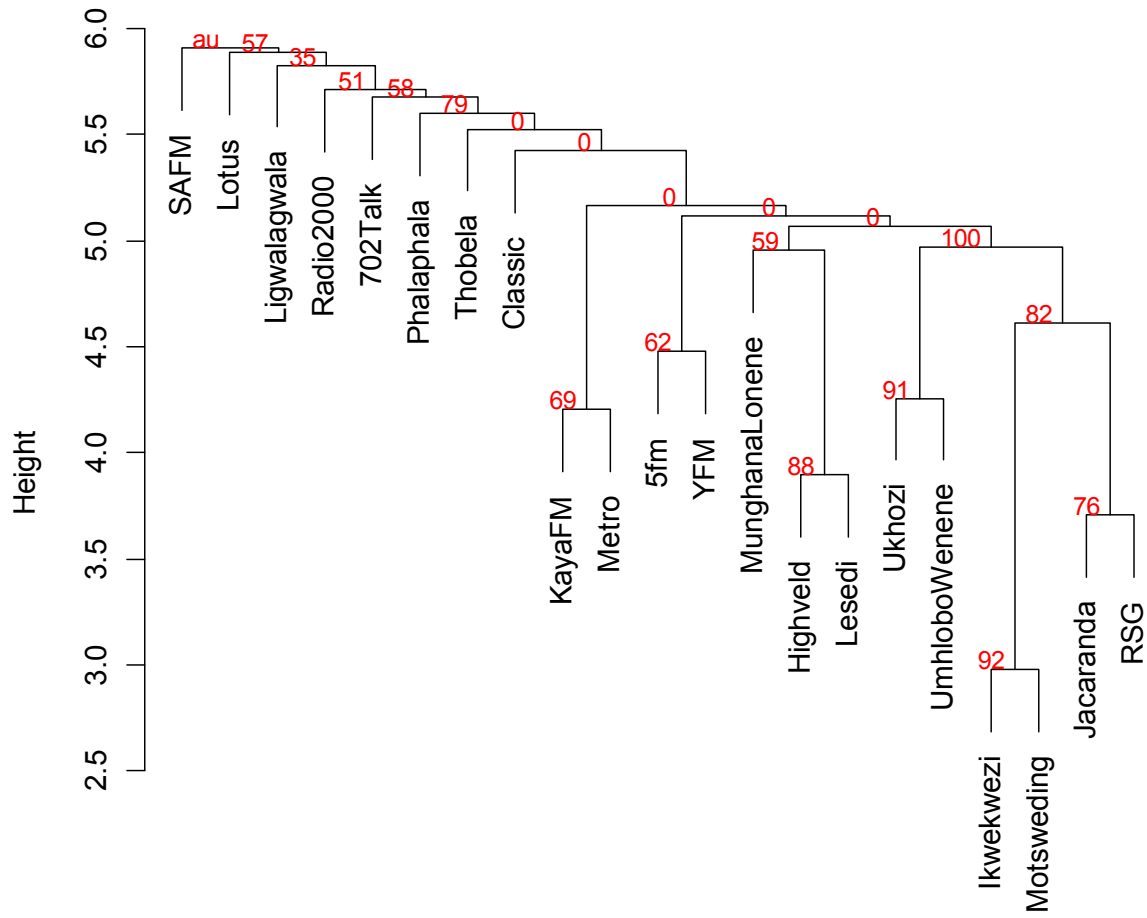
Figure 5. Gap statistic for dendrogram based on radio listener characteristics (expressed in proportions), average linkage method and statistical distance.
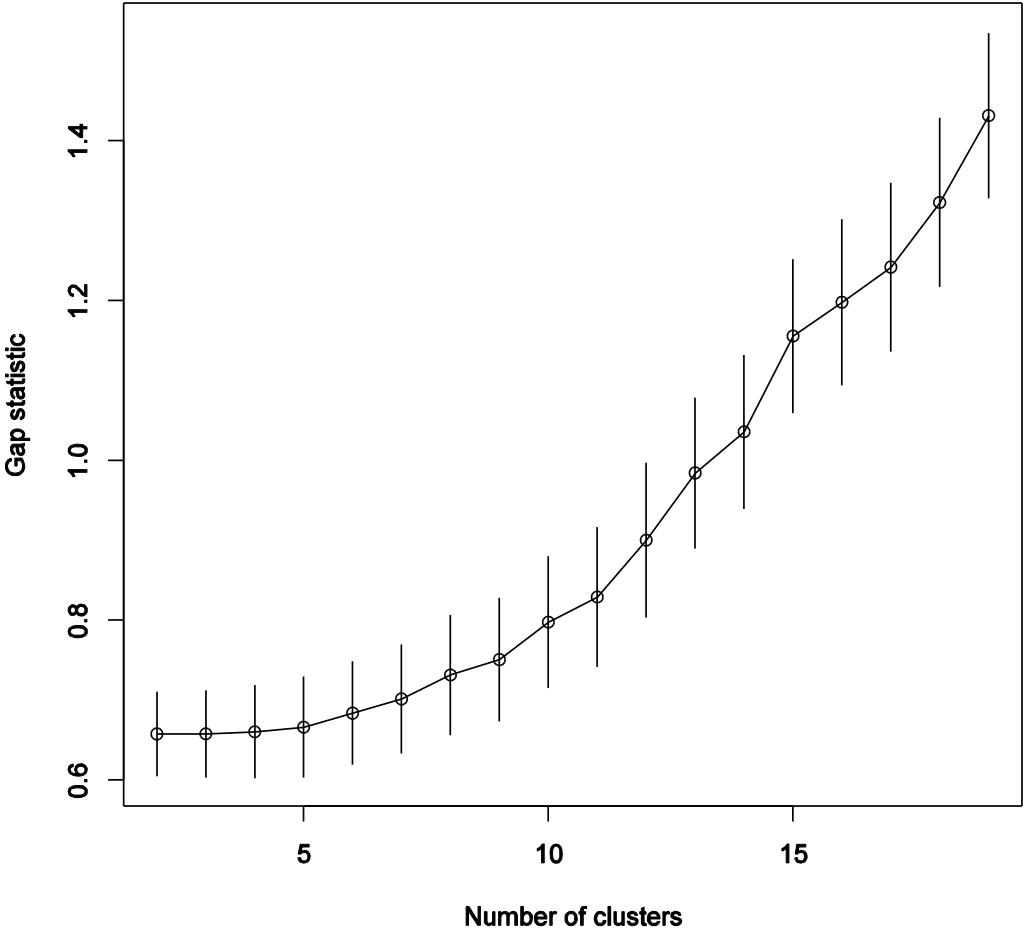
Figure 6. Dendrogram based on radio listener characteristics, using the single linkage method and statistical distance.
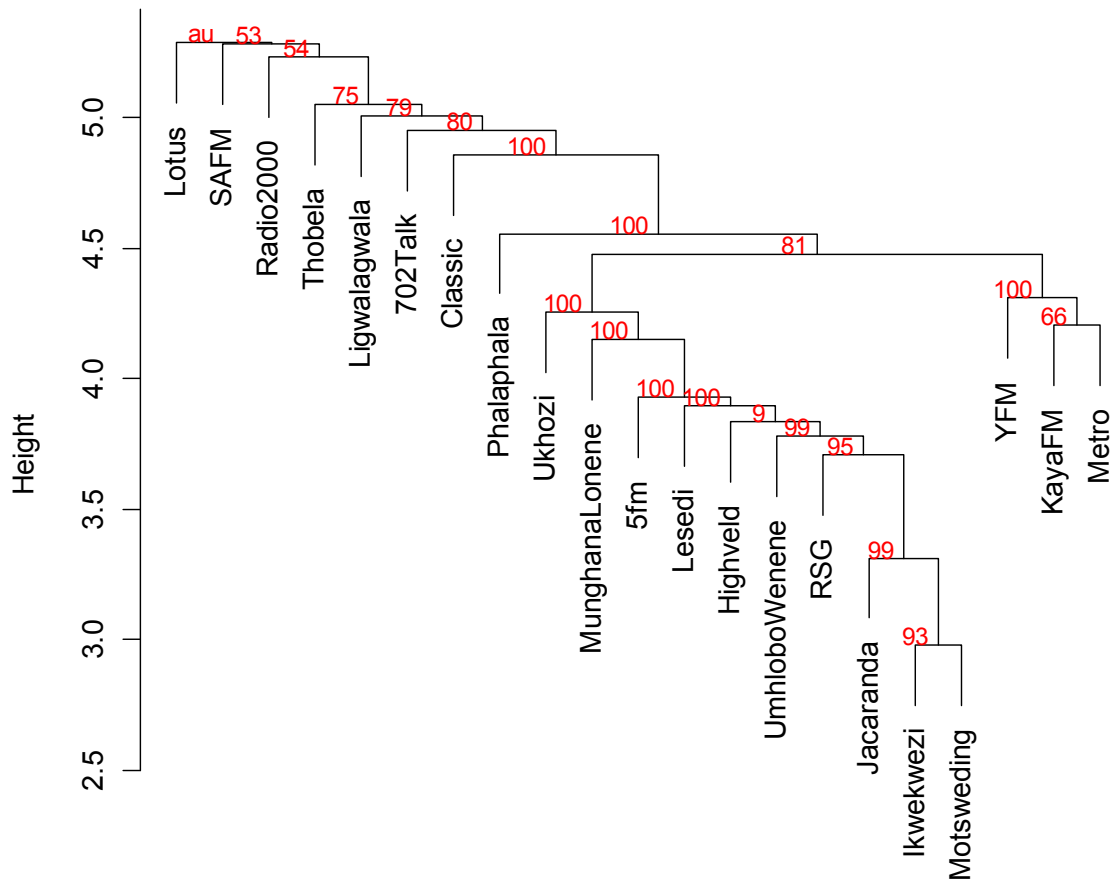
Figure 7. Gap statistic for dendrogram based on radio listener characteristics (expressed in proportions), single linkage method and statistical distance.
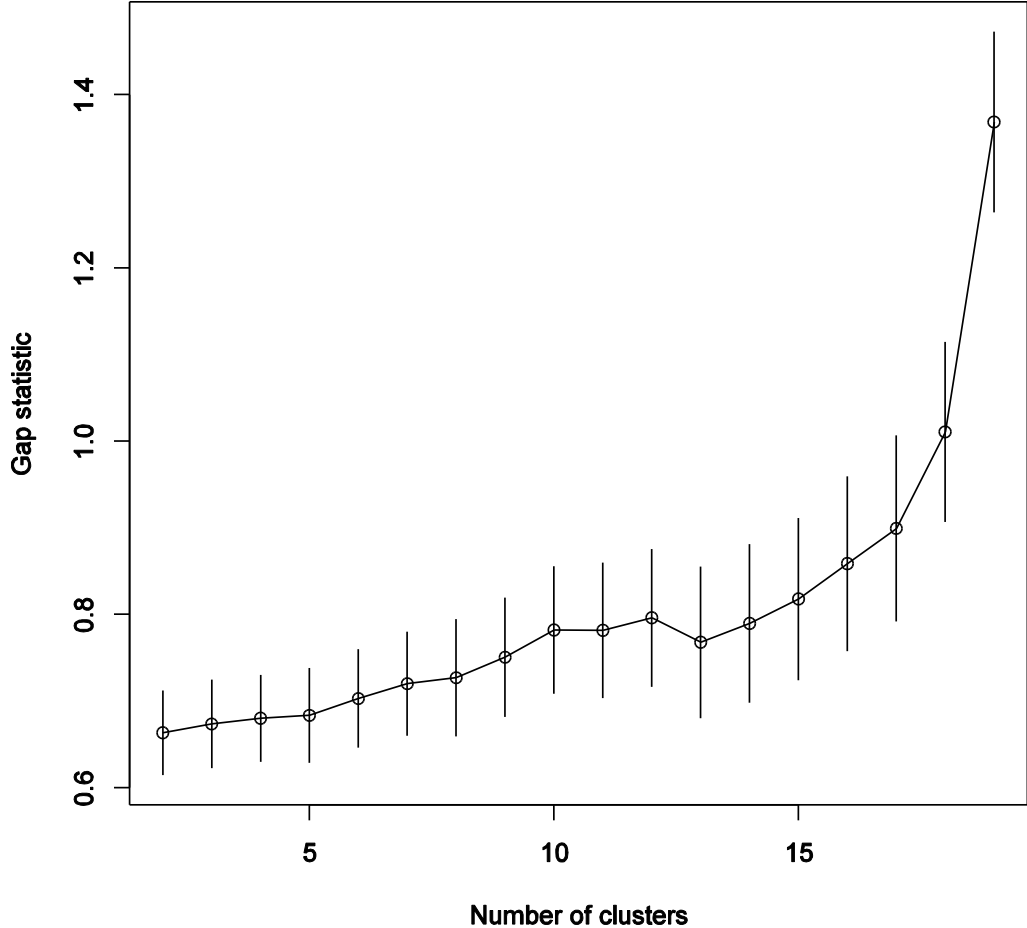
Figure 8. Dendrogram based on radio listener characteristics, using the complete linkage method and statistical distance
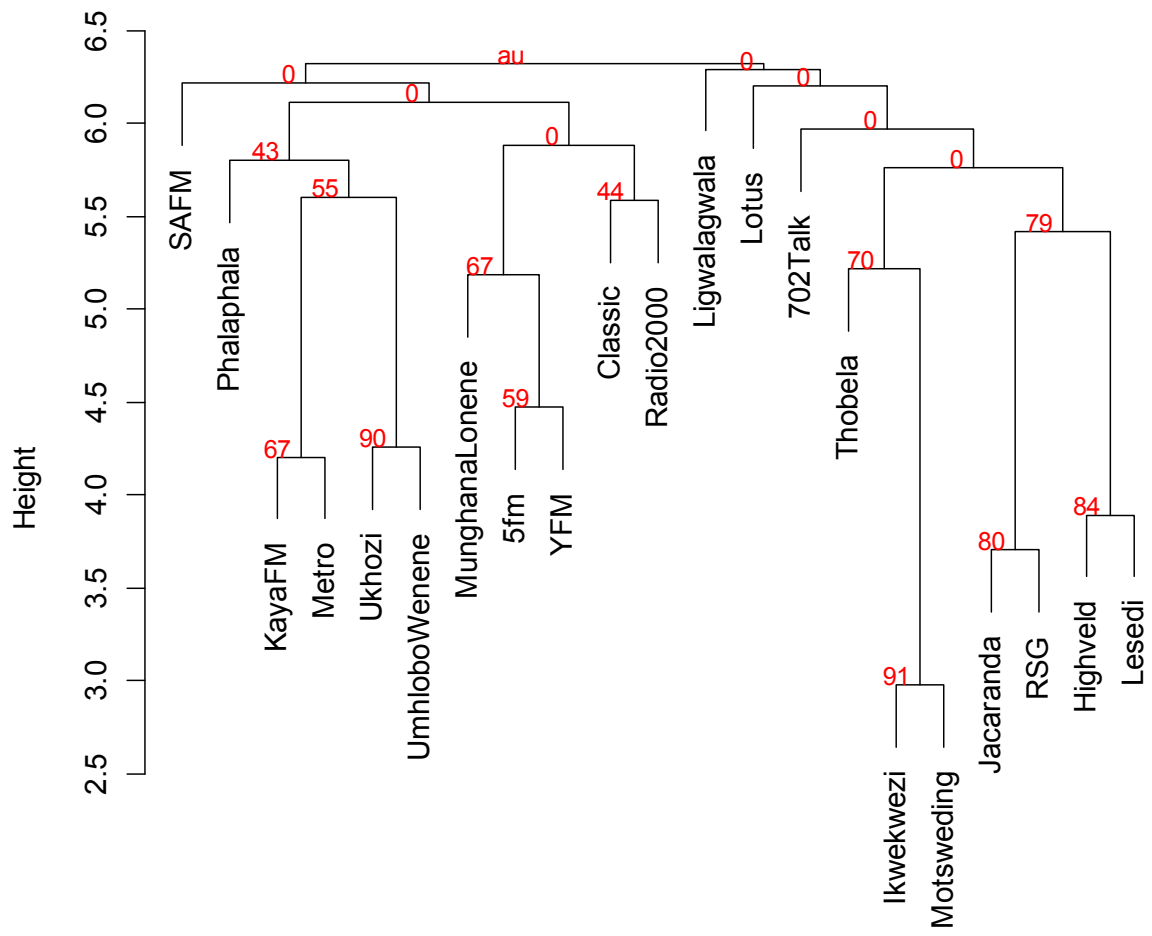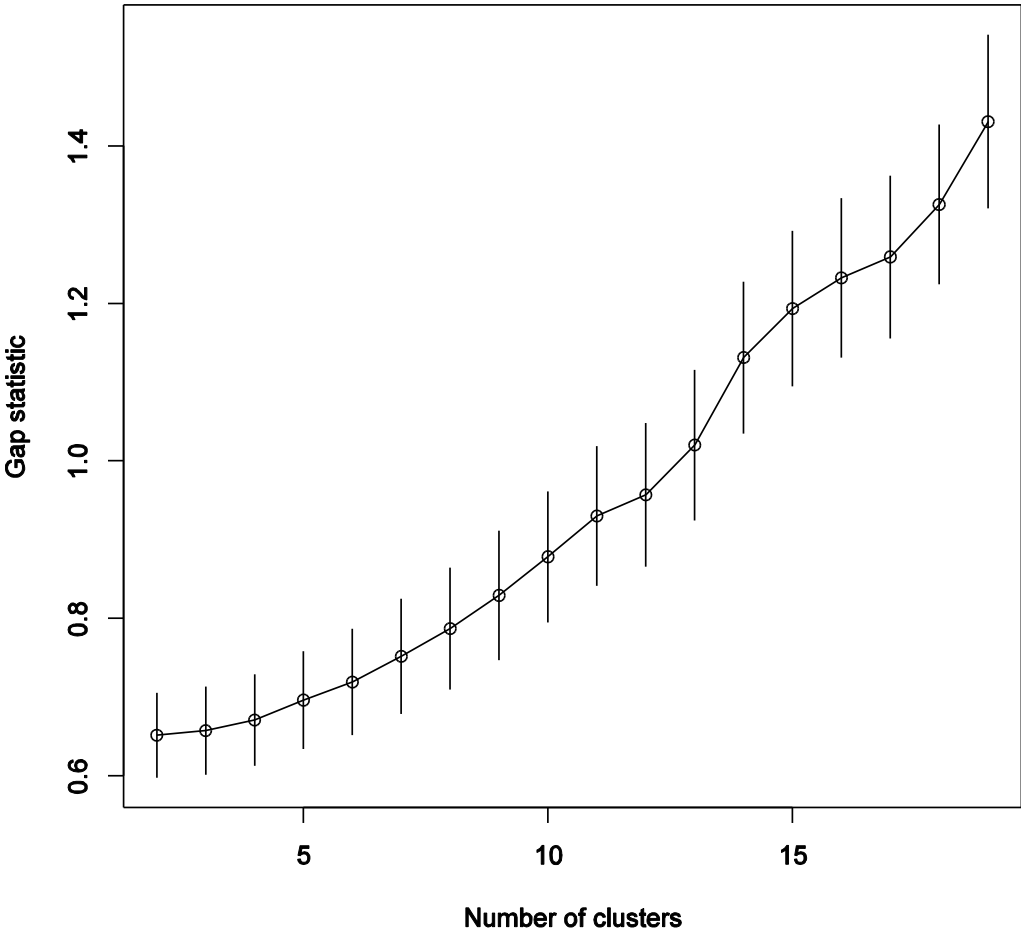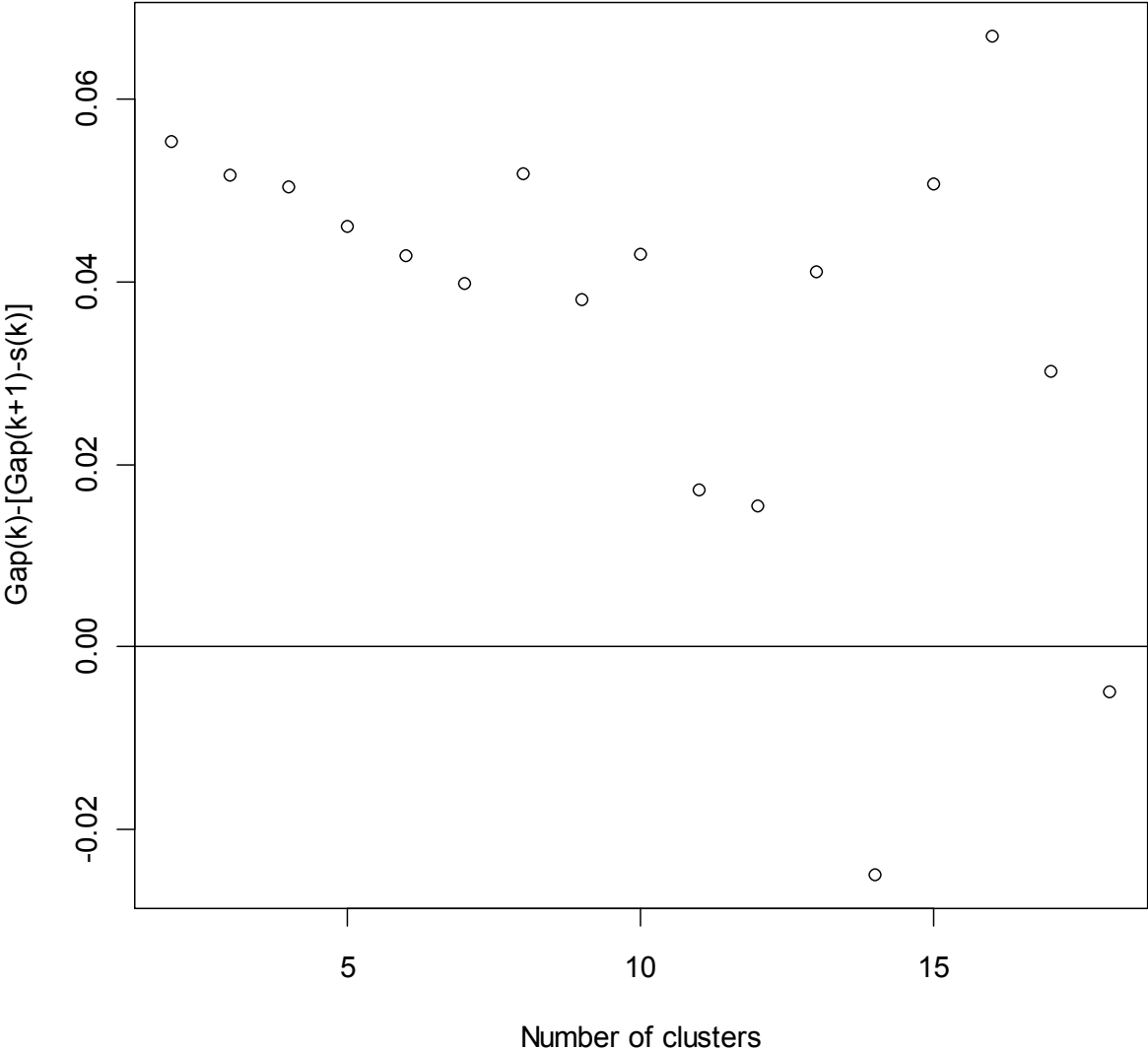
Figure 9. Gap statistic for dendrogram based on radio listener characteristics (expressed in proportions), complete linkage method and statistical distance
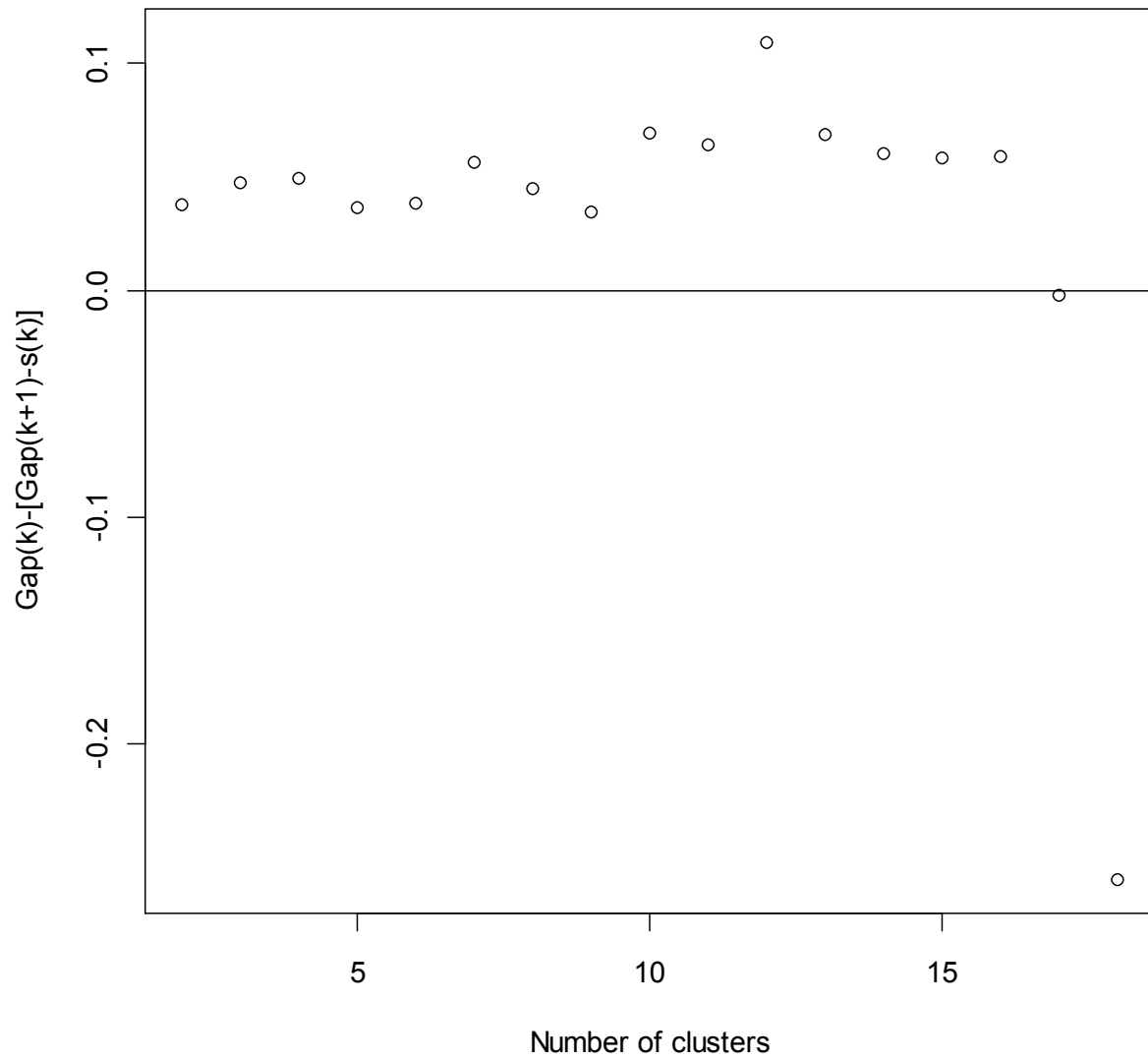
Appendix B: 1-standard deviation rule for gap statistic

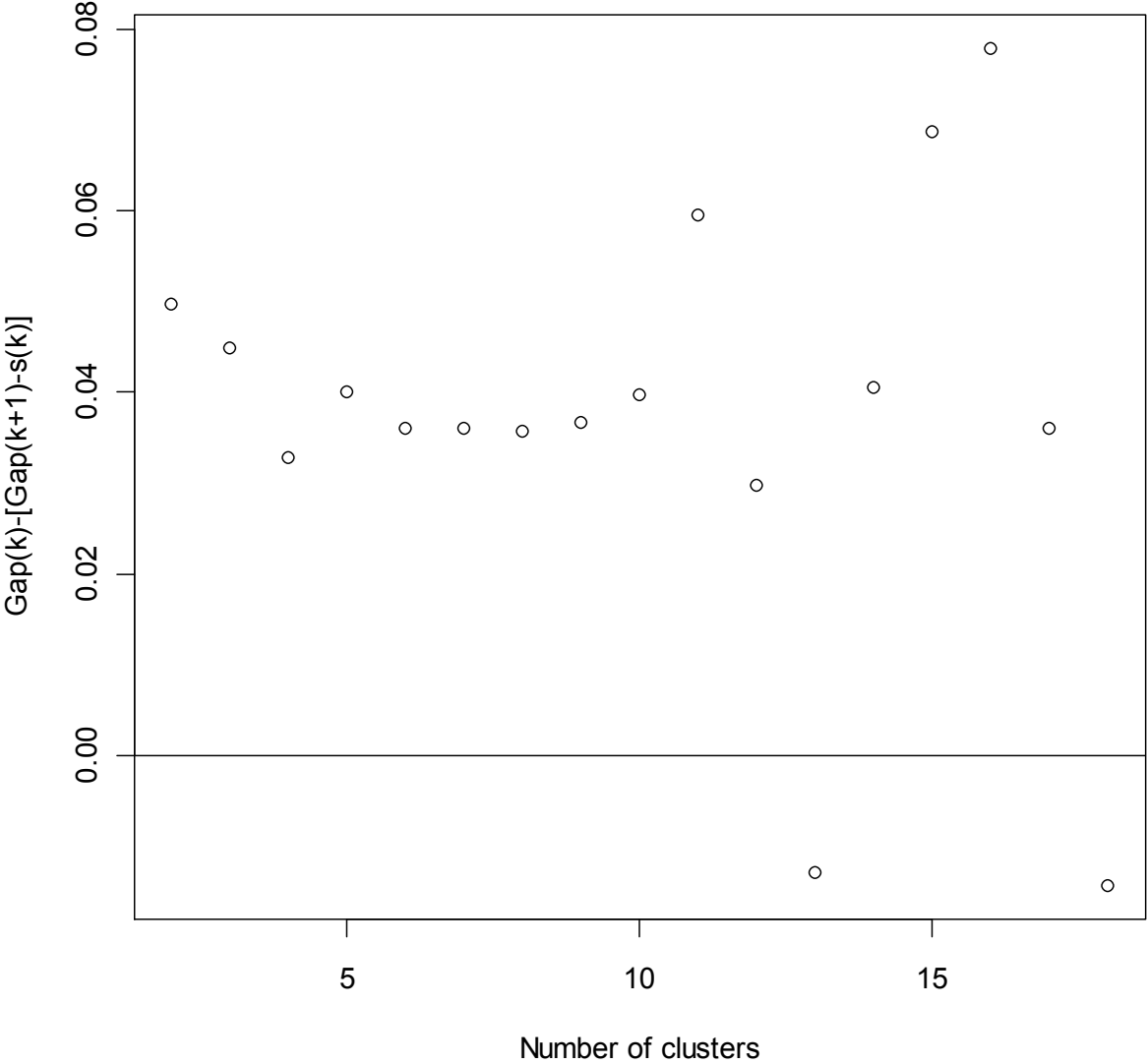B.1   <u>Average linkage method</u>

B.2   Single linkage method

Table 1. RAMS data on listener profile of Gauteng radio stations.

Listener characteristic	Description
Income group	Ten categories: LSM 1 – LSM 10
Race	Four categories: white, black, Coloured and Indian
Age	Four categories: 16-24, 25-34, 35-49, 50+
Sex	Two categories: male, female

Table 2. Relevant markets if Highveld and Kaya in the same market (radio stations in clusters with 80% or higher bootstrap probability highlighted in grey).

Average linkage method	Single linkage method	Complete linkage method
Kaya	Kaya	Kaya
Metro	Metro	Metro
5FM	5FM	5FM
YFM	YFM	YFM
Munghana Lonene (1)	Munghana Lonene	Munghana Lonene (1)
Highveld (1)	Highveld	Highveld (1)
Lesedi (1)	Lesedi	Lesedi (1)
Ukhozi (2)	Ukhozi	Ukhozi (2)
Umhlobo Wenene (2)	Umhlobo Wenene	Umhlobo Wenene (2)
Ikwekwezi (2)	Ikwekwezi	Ikwekwezi (3)
Motsweding (2)	Motsweding	Motsweding (3)
Jacaranda (2)	Jacaranda	Jacaranda (4)
RSG (2)	RSG	RSG (4)
		Thobela
		702Talk
		Lotus
		Ligwalagwala
		Radio 2000
		Classic
		Phalaphala
		SAFM

Table 3. Ranking of preferred alternative radio stations for Highveld, 702 and Kaya listeners, 2006

Rank	Highveld		702		Kaya	
	Substitute	Proportion	Substitute	Proportion	Substitute	Proportion
1	Metro	21.4	Highveld	37.1	Metro	64.9
2	5fm	17.8	Metro	21.0	Lesedi	40.3
3	Jacaranda	17.7	Kaya	16.4	YFM	38.9
4	Kaya	15.5	Jacaranda	14.7	Ukhozi	27.2
5	Lesedi	10.9	Lesedi	13.6	Motsweding	21.5
6	702	8.0	5fm	12.9	Highveld	19.7
7	RSG	7.8	SAFM	9.3	Umhlobo Wenene	11.9
8	Ukhozi	7.4	RSG	8.2	Ikwekwezi / Thobela	8.5

Appendix A: Description of radio consumer characteristics data

The RAMS data reports, for each radio station, the number of listeners (in thousands) falling into a particular age, race, sex and income group. Table A1 reports an extract of this data:

Table A1: Extract from original RAMS data showing number of listeners (in thousands) per age group and income group for selected Gauteng radio stations

Radio station	Age				Income					Total number of listeners
	16-24	25-34	35-49	50+	LSM 6	LSM 7	LSM 8	LSM 9	LSM 10	
Highveld	236	292	315	182	146	129	133	239	378	1025
702	9	34	68	104	46	18	37	48	66	215
Kaya FM	120	158	214	85	296	122	85	57	18	577

Source: South African Advertising Research Foundation (2005)

As discussed, the cluster analysis relies on proportions data, calculated from the raw data in Table A1. For example, the proportion of 16-24 listeners are calculated as the number of 16-24 listeners divided by the total number of listeners. Proportions therefore sum to 1 for each dimension (age, income, etc). Table A2 reports the proportions data calculated from Table A1:

Table A2: Proportions calculated from RAMS data, as shown in Table A1

Radio station	Age				Income				
	16-24	25-34	35-49	50+	LSM 6	LSM 7	LSM 8	LSM 9	LSM 10
Highveld	23.0	28.4	30.8	17.8	14.3	12.6	13	23.3	36.9
702	4.2	15.9	31.7	48.3	21.3	8.2	17.2	22.4	30.8
Kaya FM	20.7	27.4	37.1	14.7	51.2	21.1	14.7	9.8	3.1