Conceptual and Empirical Advances in Antitrust Market Definition With Application to South African Competition Policy

by

Willem Hendrik Boshoff

Dissertation presented for the degree of
Doctor of Philosophy (Economics) at the
University of Stellenbosch

Promoter: Prof. Stan du Plessis
Co-promoter: Dr Nicola Theron
Faculty of Economic and Management Sciences
Department of Economics

December 2011

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2011

Abstract

Delineating the relevant product and geographic market is an important first step in competition inquiries, as it permits an assessment of market power and substitutability. Critics often argue that market definition is arbitrary and increasingly unnecessary, as modern econometric models can directly predict the competitive effects of a merger or anti-competitive practice. Yet practical constraints (such as limited data) and legal considerations (such as case law precedence) continue to support a formal definition of the relevant market. Within this context, this dissertation develops three tools to improve market definition: two empirical tools for cases with limited data and one conceptual decision-making tool to elucidate important factors and risks in market definition.

The first tool for market definition involves a systematic analysis of consumer characteristics (i.e. the demographic and income profiles of consumers). Consumer characteristics can assist in defining markets as consumers with similar characteristics tend to switch to similar products following a price rise. Econometric models therefore incorporate consumer characteristics data to improve price elasticity estimates. Even though data constraints often prevent the use of econometric models, a systematic analysis of consumer characteristics can still be useful for market definition. Cluster analysis offers a statistical technique to group products on the basis of the similarity of their consumers' characteristics. A recently concluded partial radio station merger in South Africa offers a case study for the use of consumer characteristics in defining markets.

The second tool, or set of tools, for defining markets involves using tests for price co-movement. Critics argue that price tests are not appropriate for defining markets, as these tests are based on the law of one price – which tests only for price linkages and not for the ability to raise prices. Price tests, however, are complements for existing market definition tools, rather than substitutes. Critics also argue that price tests suffer from low statistical power in discriminating close and less close substitutes. But these criticisms ignore *inter alia* the role of price tests as tools for gathering information and the range of price tests with better size and power properties that are available, including new stationarity tests and autoregressive models. A recently concluded investigation in the South African dairy industry offers price data to evaluate the market definition insights of various price tests.

The third tool is conceptual in nature and involves a decision rule for defining markets. If market definition is a binary classification problem (a product is either 'in' or 'out' of the market), it faces risks of misclassification (incorrectly including or excluding a product). Analysts can manage these risks using a Bayesian decision rule that balances (1) the weight of evidence in favour of and against substitutability, (2) prior probabilities determined by previous cases and economic research, and (3) the loss function of the decision maker. The market definition approach adopted by the South African Competition Tribunal in the Primedia / Kaya FM merger investigation offers a useful case study to illustrate the implementation of such a rule in practice.

Opsomming

Mededingingsake neem gewoonlik ’n aanvang met die afbakening van die relevante produk- en geografiese mark. Die markdefinisie-proses werp dikwels lig op markmag en substitusie-moontlikhede, en ondersteun dus die beoordeling van ’n mededingingsaak. Markdefinisie word egter deur kritici as arbitrêr en selfs onnodig geag, veral aangesien ekonometriese modelle die uitwerking van ’n samesmelting of ’n teen-mededingende praktyk op mededinging direk kan voorspel. Tog verkies praktisyns steeds om markte formeel af te baken op grond van sowel praktiese oorwegings (insluitend databeperkings wat ekonometriese modellering bemoeilik) as regsoorwegings (insluitend die rol van presedentereg). Hierdie proefskrif ontwikkel dus drie hulpmiddels vir die definisie van markte: twee empiriese hulpmiddels vir gevalle waar data beperk is sowel as ’n denkhulpmiddel om o.a. risiko’s rondom markdefinisie te bestuur.

Die eerste hulpmiddel vir die definisie van markte behels die sistematiese analise van verbruikerseienskappe, insluitend die demografiese en inkomste-profiel van verbruikers. Verbruikerseienskappe werp lig op substitusie, aangesien soortgelyke verbruikers neig om na soortgelyke produkte te verwissel na aanleiding van ’n prysstyging. Ekonometriese modelle maak derhalwe van data omtrent verbruikerseienskappe gebruik om beramings van pryselastisiteit te verbeter. Hoewel databeperkings dikwels ekonometriese modellering beperk, kan verbruikerseienskappe op sigself steeds nuttig wees vir die afbakening van die mark. Trosanalise bied ’n statistiese metode vir ’n stelselmatige ondersoek van verbruikerseienskappe vir markdefinisie, deurdat dit produkte op grond van gelyksoortige verbruikerseienskappe groepeer. ’n Onlangse ondersoek in Suid-Afrika rakende die gedeeltelike samesmelting van Primedia and Kaya FM radiostasies bied data om die gebruik van trosanalise en verbruikerseienskappe vir markdefinisie-doeleindes te illustreer.

Die tweede hulpmiddel vir markdefinisie behels statistiese toetse vir verwantskappe tussen prystydreekse van verskillende produkte of streke. Hierdie prystoetse is gebaseer op die wet van een prys en beklemtoon prysverwantskappe eerder as die vermoë om pryse te verhoog (wat die uiteindelike fokus in mededingingsbeleid is). Hierdie klem verminder egter nie noodwendig die insigte wat prystoetse bied nie, aangesien markdefinisie dikwels ’n omvattende analise verg. Prystoetse se statistiese onderskeidingsvermoë word ook dikwels deur kritici as swak beskryf. Hierdie tegniese kritiek beskou prystoetse as eng-gedefinieerde hipotesetoetse eerder as hulpmiddels vir die verkenning van substitusiepatrone. Voorts ignoreer hierdie tegniese kritiek ’n verskeidenheid nuwe prystoetse met beter onderskeidingsvermoë, insluitend nuwe toetse vir stasionêriteit en nuwe autoregressiewe modelle. ’n

Onlangse mededingingsondersoek in die Suid-Afrikaanse melkindustrie verskaf prysdata om die verrigting van verskillende prystoetse vir geografiese markdefinisie te ondersoek.

Die derde hulpmiddel vir die definisie van markte behels ’n besluitnemingsreël. Hiervolgens word markdefinisie as ’n binêre klassifikasieprobleem beskou, waar ’n produk of streek ‘binne’ of ‘buite’ die mark geplaas moet word. Gegewe dat hierdie klassifikasie onder toestande van onsekerheid geskied, is markdefinisie blootgestel aan risiko’s van wanklassifikasie. Praktisyns kan hierdie risiko’s bestuur deur gebruik te maak van ’n Bayesiaanse besluitnemingsreël. Sodanige reël balanseer (1) die gewig van getuienis ten gunste van en teen substitusie, (2) *a priori* waarskynlikhede soos bepaal deur vorige mededingingsake en akademiese navorsing, en (3) die verliesfunksie van die besluitnemer. Die benadering van die Suid-Afrikaanse Mededingingstribunaal in die saak rakende die gedeeltelike samesmelting van Primedia en Kaya FM bied ’n nuttige gevallestudie om hierdie beginsels te demonstreer.

Acknowledgements

I thank Prof. Stan du Plessis and Dr Nicola Theron for their careful attention and thoughtful comments on the many draft papers that preceded this dissertation.

I also thank Prof. Andrie Schoombee, chairperson of the Department of Economics, for his generous management of my teaching and other responsibilities at the Department over the past two years.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction and Research Questions

Competition law investigations usually commence with a definition of the relevant product and geographic market. The relevant market provides a first evaluation of competitive conditions and allows for the calculation of market shares, which aids in the assessment of firms' market power. Given its implications for assessing market power, the market definition in a competition case is frequently contested. Observers have long argued that market definition is arbitrary and, more recently, econometricians have highlighted that modern econometric methods are capable of directly estimating market power and competitive effects. Despite these criticisms, practitioners have retained the market definition as the first step in most competition investigations. Practitioners note that practical data constraints prevent large-scale econometric modelling in most cases and that legal issues – including case law precedence and preferences for legal certainty – favour retaining the market definition as a first step.

This dissertation focuses on two important contributions for improving market definitions. Firstly, the dissertation develops empirical tools with limited data requirements that are useful for defining markets. The tools are cluster analyses applied to data on consumer characteristics and tests for time-series relationships applied to price data. Secondly, the dissertation employs statistical decision theory to develop a conceptual framework for defining markets. The conceptual framework is useful to practitioners, as it highlights the roles of three factors involved in all such cases: case evidence, prior probabilities (informed by previous cases and economic research) and the loss functions of the various parties to the case.

The conceptual and empirical tools developed in this dissertation are influenced by a range of developments in economics, econometrics and law that have shaped how markets are defined in competition law investigations. This chapter provides an overview of these developments and their ultimate linkages to the three tools developed in this dissertation.

1.1 Introduction

Ever since Adam Smith's first description of the 'invisible hand', economists have been interested in the unconscious coordination of productive activity via markets (Smith, 1776). Price is determined in the 'market', a notional (and sometimes a physical) space where a set of buyers and a set of sellers of a particular commodity or service repeatedly interact. Right from the start of economics as a discipline then, the concept of a market has held a prominent position in the vocabulary of economists. In law too, the market concept has gradually acquired important status over the past one hundred years. The rise of antitrust policy in the US since the late nineteenth century and its spread to other jurisdictions has gradually stimulated the need for greater precision in evaluating market power and, hence, market boundaries.

Despite its academic and legal prominence, the market and in particular, the extent of the market, did not receive explicit research attention until comparatively recently. It was only in the early 1980s that economists started responding to the problem of 'market definition' in competition law investigations, i.e. the problem of deciding on the extent or boundaries of the market (see Stigler and Sherwin (1985) and others in Massey (2000)). Since then, economists working on competition policy and industrial organisation have developed guidelines for defining markets based on economic reasoning. Economists also started developing quantitative tools for market definition in the 1980s (for recent surveys see Baker and Bresnahan (2008), Coate and Fischer (2008) and Carlton (2007)).

Despite these advances in market definition, the practice of formally delineating market boundaries remains controversial. The controversy is not limited to academics, as indicated by the recent debate among regulators and practitioners on the market definition implications of an increasingly effects-based approach and of the revised US merger guidelines (see, for example, Arezzo (2008), Keyte and Stoll (2004) and Markovits (2002)). This dissertation contributes to the debate among academics and practitioners and aims to develop further conceptual and empirical tools to improve market definition.

This chapter provides the conceptual and historical context for the market definition tools developed in later chapters. The chapter starts with an exploration of the market concept in its various forms in US, EU, UK and South African competition policy. The chapter then considers whether market definition is still relevant to modern competition policy, after which it identifies the three research questions, which receive closer attention in subsequent chapters.

1.2 The market concept in competition policy

The term 'market' is often used colloquially as a reference to capitalism, and even economists tend to use the concept loosely when referring to a process of spontaneous exchange. But the rise of policy intervention in a variety of markets and an increased understanding of the institutional structure of production have led academic economists to view a 'market' as a well-defined space with clear boundaries. For example, in institutional economics, Coase (1937) and later Williamson (1975) focused on how transaction costs determine boundaries between firms and their input and output markets. In regulatory policy, interventions in telecommunications and utility markets gave rise to a need for the explicit delineation of market boundaries (see Theron and Boshoff (2006) for a recent South African application). But it was especially in competition law investigations that the exact definition of the relevant market became an important topic.

As discussed below, competition policy was developed with the aim of constraining market power, which is commonly defined as "the ability of a firm to raise the prices of its products above the competitive level" (Davis and Garcés, 2010: 162). In early competition policy, this ability of a firm to raise prices was thought to be strongly linked to its size – the larger the firm, the greater its market power. To measure size, competition authorities relied on the market shares of firms. But the calculation of market shares necessarily requires a definition of the relevant market. Therefore, during the 1970s and 1980s US and EU courts *required* the exact definition of the relevant market. In the US, this happened in the famous Brown Shoe case, and in the EU, after the merger case involving Europemballage Corporation and Continental Can Company Inc (Davis and Garcés, 2010: 161).

Market definition became increasingly sophisticated in the 1960s and 1970s, stimulating demand for the quantitative toolkit and the theoretical models of economists. In fact, one could argue that it was specifically in the field of market definition that economics started its gradual ascent in competition policy. However, the increased use of economists did not necessarily render market definition straightforward. While economics can offer significant insights into markets and their processes, the market concept in economics cannot be applied directly to competition policy. There are remarkable differences between market concepts in the various subdisciplines of economics, and these differences are particularly relevant to competition policy, where the exact definition of the market is frequently determinative.

Competition policy focuses on the artificial creation or abuse of market power so that a market is defined as the smallest product and geographic space that is worth monopolising, i.e. it will include all substitutes that constrain the market power of the firm being investigated (see subsequent sections). A strategic market, in turn, is the smallest possible space in which to be a viable competitor, where 'viable' refers to profitability and 'smallest' refers to strategic necessity (Kay, 1990). For example, while serving a global market may be a strategic option, local niches may be more attractive so that the strategic market is local despite the option of being a global competitor. An economic or trading market is yet another concept, referring to the smallest space in which the firm "will be forced to charge all consumers exactly the same price for the same good" (Geroski, 1998: 691). This definition is closely linked to the so-called law of one price, which receives attention in Chapter 3. Clearly, the set of substitutes that meet the criterion for inclusion in a market for competition policy need not be the same set of products that are included in the strategic or economic market. This variety in market concepts can be a source of confusion in competition investigations. For example, testimonies of businesspeople and economists before the competition authorities can be based on quite different market concepts (Geroski, 1998).

The difference between a market in a competition law investigation and a strategic market is also recognised in the management strategy literature. This literature distinguishes between a 'natural' market and a so-called 'enacted' market, and the market concept in competition policy is closely related to that of a natural market (Brooks 1995). The natural market is a collective construct which assumes that firms compete in a common space that can be identified *independently* of the views of a particular firm. This definition of a market accords with the structure-conduct-performance (SCP) paradigm (Brooks, 1995: 537, emphasis added):

> "Studies of the performance implications of structural relationships amongst suppliers must be conducted in the context of markets … since performance effects are dependent on *the interactions between suppliers and customers and on the competition amongst suppliers seeking to serve the same of customers*".

However, businesspeople may have an 'enacted' view of the market, where the market space is an evolutionary, firm-specific construct. In this view, the market space differs depending on the perspective of the particular organisational actor involved. This market construct is consistent with the views of Hayek, who highlighted the emergent nature of economic activity: an individual firm rarely has comprehensive knowledge of the entire market and the various players, and advances its localised knowledge of demand and supply in an idiosyncratic, piecemeal fashion. This inevitably results in different views of 'the' market among firms that are grouped into the same 'natural' market.

Competition policy has not been blind to the fact that its concept of the market as a common space may over-simplify complex commercial reality, which may be important in a competition investigation. The debate surrounding an effects-based approach to competition policy suggests that practitioners and academics increasingly recognise that the size of a firm in a common space is not necessarily the best indicator of its market power or the effects of its conduct (see Kovacic and Shapiro (2000) for a US perspective; Gual, Hellwig et al. (2005) for a European perspective). In the South African context, Corbett, Das Nair and Roberts (2010) discuss the role of market definition and other factors in the assessment of market power, while Theron and Boshoff (2011) consider the role of traditional analysis (based on market definition and the calculation of market shares) under an effects-based approach to anti-competitive conduct.

If market definition is important and refers to a common space that includes all competitors that constrain the market power of the firm under investigation, it is useful to consider how different jurisdictions (the US, EU and South Africa) approach the market definition exercise. In particular, it is important to consider how the different jurisdictions measure competitive constraints.

1.2.1 Developments in the US

The Sherman Act, passed in 1890, and later the Clayton Act, passed in 1914, sought to prevent horizontal and vertical restraints, abuse of dominance practices, and price discrimination in the US. In addition, the Clayton Act also introduced the world's first merger control regime. More important, this Act also explicitly required the definition of both a product and a geographic market (see the summary in Blair and Kaserman (2009: 61)). Market definition therefore had already received explicit attention in the early parts of the twentieth century.

Over time, and especially following the rise of economics in competition policy since the 1980s, the US Department of Justice and the Federal Trade Commission (FTC) developed guidelines on how market definition is to be performed. Specifically, these agencies released the so-called Horizontal Merger Guidelines, in which they noted explicitly the type of tests and forms of evidence necessary for antitrust market definition (United States Department of Justice and Federal Trade Commission, 1992; 2010).

The Guidelines introduced the hypothetical monopolist (HM) test for market definition in a competition policy setting. The HM test views a market as that product and geographic space that can *potentially* be monopolised by the firm(s) being investigated (Geroski, 1998). The emphasis is on identifying those firms and regions which act as competitive constraints on the firm, preventing it from using its power to raise prices profitably. In fact, the HM test is frequently phrased in terms of a thought experiment, in which the competition analyst defines the relevant geographic market by considering whether the firm under investigation is capable of maintaining a price increase of 5%-10% for a twelve-month period (for example) without a reduction in profits (referred to as a 'small but significant non-transitory increase in price' (SSNIP)). The SSNIP test starts with only the geographic area in which the firm under investigation is operating. If the firm's profits are ultimately adversely affected by the price increase, the geographic market is too narrow. Consequently, a broader geographic market can be defined by including that region from which competition is most likely to originate following the price increase. The thought experiment is repeated and other regions are added until a broad enough geographic market has been defined in which the firm under investigation could raise prices on a profitable and sustainable basis. A similar exercise can be carried out to delineate the product market.

The SSNIP test is a thought experiment and, in practice, empirical measures are required to operationalise the test. Specifically, the Guidelines state that measures of demand-side substitutability are the appropriate empirical measures for market definition (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."

Measures of demand-side substitutability usually employed in the US include direct econometric estimates of own- and cross-price elasticity of demand or, as is usually the case, indirect estimates inferred from price relationships, product flows, consumer surveys, industry expert opinions and other qualitative information (see Baker (2007) for a recent summary).

Supply-side substitutability is also important to market power assessment, but receives less attention during the market definition stage. For supply-side substitutability to influence market definition, substitution must be "easy, rapid and feasible" (Motta, 2004: 105), which are stringent conditions in a variety of market settings. Nevertheless, there are conditions where even the threat of entry will constrain price (see, for example, the work on contestability by Baumol, Panzar and Willig (1982)). Notwithstanding this strand of literature, the focus in market definition has remained on demand-side substitutability, and most jurisdictions assign a secondary role to supply-side substitution (Davis and Garcés, 2010).

1.2.2 Developments in the EU and UK

Competition policy in Europe did not receive extensive attention until the Treaty of Rome in 1957. Articles 81 and 82 of the Treaty established the EU competition framework, but with the exception of the UK and Germany, serious competition law in member states only took off from the early 1980s. The EU borrowed from (although did not replicate) the American system to develop its competition framework (Neven, 2006). The EU adopted a similar approach, with the Articles of the Treaty (like the American Sherman or Clayton Acts) describing the broad contours of competition policy and the need for market definition. Subsequent Commission Notices (like the Horizontal Merger Guidelines) elaborated on the guiding principles for defining product and geographic markets. In the EU, as is the case in the US, competition policy relies exclusively on demand-side substitutability for market definition (Davis and Garcés, 2010: 165).

The UK operates a competition policy regime separate from that of the EU despite its EU membership. Nevertheless, the UK competition regime follows that of the EU in assigning a primary role to demand-side substitutability in market definition. The UK explicitly adopted the hypothetical monopolist test and the SSNIP thought experiment for market definition. In the 2010 Merger Assessment Guidelines, the Office of Fair Trading (OFT) and the UK Competition Commission (CC) note the following (Office of Fair Trading and Competition Commission, 2010: 30):

> "In identifying the relevant product market, the Authorities will pay particular regard to demand-side factors (the behaviour of customers and its effects). However, they may also consider supply-side factors (the capabilities and reactions of suppliers in the short term) and other market characteristics."

In both the EU and the UK, the type of empirical evidence to measure such demand-side substitutability is quite similar to that of the US, and includes direct and indirect estimates of price elasticity and diversion ratios, including econometric elasticity estimates, tests of price relationships, information on product characteristics, consumer surveys and internal business documents (Office of Fair Trading and Competition Commission, 2010: 32).

1.2.3 The South African experience

Market definition enjoys an important position in South African competition policy. South African policy makers consulted widely when drafting the competition policy regime for post-Apartheid South Africa[1], drawing on EU, US and Canadian competition policy (Organisation for Economic Co-operation and Development, 2003: 21). The subsequent Competition Act (Act No 89) of 1998 introduced *inter alia* market shares to assess market power (Sutherland and Kemp, 2000), which implies an important position for market definition.

The 1998 Act specifies in Section 7 the following market share conditions for assessing whether a firm is dominant, i.e. whether it has market power (Republic of South Africa, 1998):

- If its market share is greater or equal to 45 percent of the market, the firm is considered dominant.
- If its market share is greater than 35 percent but less than 45 percent, the firm should prove that it does not possess market power. Otherwise the firm is also assumed dominant.

[1] Apartheid South Africa did have a competition policy regime, but the previous Competition Board did not have the same investigative and punitive powers as those allocated to South African competition authorities under the new regime (Roberts, 2004).

- If its market share is below 35 percent but the firm is deemed to have market power, the firm is assumed dominant.

The market share conditions for inferring market power were met with fierce criticism from some economists, notably Reekie (1999), who argued that the market share thresholds are arbitrary and reflect an underlying SCP paradigm. The Reekie criticism of the Act links with the broader academic and policy debate of the nineties concerning the relationship between high concentration levels in the South African economy and their implication for economic performance. Fourie and Smith (1993) and Fourie (1996) argued that concentration was detrimental to economic performance, while others such as Reekie (1999) and Leach (1992) concluded differently. Fourie and Smith (1998; 1999) highlight the main fault lines in the debate, which has subsequently continued – recent research by Fedderke and Naumann (2009), Fedderke and Szalontai (2009) and Fedderke and Simbanegavi (2008) support the earlier Fourie and Smith findings[2], whereas Du Plessis and Gilbert (2007; 2008) and Edwards and Van de Winkel (2005) find evidence to the contrary. Therefore, despite compelling evidence of more competitive conditions in and improved performance of the South African economy following the liberalisation efforts of the 1990s (Frankel, Smit and Sturzenegger, 2008), disagreement remains on the impact of concentration on performance in key sectors. Despite this disagreement, the prominence of concentration in the South African policy debate implies an important position for market definition.

The SA Competition Commission indicated early on that it views the SCP paradigm more as a useful organising framework and less as a suitable economic model of competition (see the discussion in Theron (2001)). The Commission highlighted that an SCP organising framework does not necessarily ignore the backward linkages from performance to concentration, as opponents of the SCP paradigm often argue. This is consistent with Smit (1999), who argues that the SCP paradigm provides a useful taxonomy for academic research on competition policy, even if concentration-profit relationships are more complex than the unidirectional form suggested by the SCP paradigm.

The discussion above suggests that competition policy practice accords market definition an important position in South Africa, as elsewhere. However, the SA Competition Act does not provide formal guidelines as to how markets are to be defined. Reekie (1999: 269) criticised the lack of a proper economic foundation for the relevant market in the Act, highlighting the need for clarity on whether demand-side or supply-side substitutability will drive market definition:

[2] Although these authors find a positive relationship between concentration and mark-ups in South Africa, they are careful to point out that the relationship is unlikely to be unidirectional.

> "The correct approach is not to opt for the narrowest or broadest market definition, but rather to examine the most appropriate definition of the relevant market for the case at hand. Given the position of adversarial advocacy often adopted by representing lawyers … it becomes ever more important that economists determine the appropriate emphasis to be laid on either the supply or demand side of the market, and in turn the appropriate level of aggregation."

In the absence of formal guidelines in the 1998 Act, practitioners have adopted a market definition approach consistent with the approach in other jurisdictions, and rely heavily on US-based conceptual tests, including the SSNIP test. Theron (2001) notes that market definition in the years immediately following the adoption of the Act focused mostly on demand-side substitutability. Nevertheless, the Act does not require market definition to be based exclusively on demand-side substitutability and both demand and supply substitution have been important in market definition (Corbett et al., 2010: 7).

Practitioners do not see the lack of formal guidelines with regard to type of evidence (demand- or supply-side) for market definition as necessarily problematic. Even if supply-side substitutability is not always considered in the market definition step, it receives attention elsewhere. SA competition practitioners, consistent with their counterparts in other jurisdictions, hold that the definition of the relevant market is only a means to the end of assessing market power (Corbett et al., 2010). The Act favours such an approach, specifying that the enhancement of market power depends on a range of competitive constraints that includes but is not limited to demand-side substitution. This suggests that demand-side substitution can be the main focus in market definition, as other factors relevant to conclusions on market power are also considered in the subsequent stages of an investigation. These factors are described in Section 12A(2) of the Act and include supply-side substitutability (in the form of an assessment of barriers to entry) as well as a range of other factors such as concentration and countervailing power (see Reekie (1999: 281-282) for an overview).

The above overview of market concepts in various jurisdictions suggests that market definition is a ubiquitous first step in competition investigations. But some have questioned the relevance of market definition to modern competition policy: these critics argue that econometric IO models allow analysts to directly assess competitive effects without the need to engage in complex market definition. While this is an important criticism given the increased focus on effects rather than form, the following section shows that market definition can provide important substitution information relevant to an analysis of effects. The section distinguishes between mergers and abuse of dominance investigations, given that the two types of investigation differ in aims – with implications for market definition.

1.3 Is market definition relevant to modern competition policy?

An assessment of the relevance of market definition should be sensitive to the type of competition investigation involved. The practice of defining the relevant market first developed in the area of merger investigations. A merger investigation traditionally involves a prospective analysis of the power likely to be created by the merger in the relevant market, and market definition featured centrally in the measurement of such market power. As the competition policy literature developed, the likely effects of a proposed merger (split between so-called ‘unilateral’ and ‘coordinated’ effects) have received increasing attention. Nevertheless, market definition retained its importance, as power in the relevant market is often seen as informative in an analysis of merger effects; as noted earlier, practitioners continue to view the SCP paradigm, even if only implicitly, as a useful framework for competition analysis.

As discussed in greater detail in the following section, the retention of market definition in merger assessment is strongly influenced by a preference for an eclectic empirical strategy. Practitioners prefer to use a range of models and tools and a variety of qualitative and quantitative evidence rather than to rely on a single encompassing model. Market definition is a useful first step to frame such a broad-based analysis. In contrast, some specialists in IO insist that a fully-specified econometric model is best placed to predict the effects of a merger. In addition, critics have argued that market definition is in any event an inefficient way of assessing market power, as it is possible to directly measure market power using these models (original work is by Baker and Bresnahan (1985; 1988); Davis and Garcés (2010) offer a more recent discussion). While practitioners have taken note of the empirical criticisms of market definition and increasingly employ advanced econometric techniques in merger assessment, market definition remains the preferred first step. A following section considers this choice in greater detail.

Apart from merger evaluation, investigations related to anti-competitive conduct, including abuse of dominance and vertical restraints, also commence with market definition. Market definition in these investigations usually follows the principles of market definition for merger cases. However, in recent years, market definition in abuse of dominance investigations has received particular attention, driven by the shift towards a so-called ‘economics-based’ or ‘effects-based’ approach to abuse of dominance cases[3].

Even if such an effects-based approach has not yet been fully or formally implemented, competition cases now focus much less on traditional form-based analysis and more on the economic effects of particular conduct (Kovacic and Shapiro, 2000; Gual et al., 2005; Roeller and Stehmann, 2005). In step with this

[3] This follows an earlier shift to an effects-based approach to vertical restraints. Theron and Boshoff (2011) summarise these developments and their relation to South African competition policy.

development, some critics have argued that market definition is less useful under an effects-based approach: dominance inferred from market share is argued to be an inappropriate measure of market power and, even if it is, there is not necessarily a causal link between dominance and the competitive effects of particular conduct (see, for example, Niels and Jenkins (2005), and the discussion in Arezzo (2008) and Fisher (2007)).

Proponents of market definition and also legal scholars opposed to breaking with established legal precedent have held that market definition for abuse of dominance cases should be retained, as it may provide a useful screen to economise on the efforts of the courts. These proponents concede that market definition is, at best, a very rough proxy for actual market power, but that it is nevertheless a useful screen: if the firm under investigation is not dominant in the relevant market (according to market share thresholds) it is less likely that that firm's conduct could have significant anti-competitive effects (Carlton, 2007). Competition authorities appear to support this view of market definition as a useful screen in abuse of dominance investigations and have retained market definition as the first step in preventing the abuse of market dominance (Office of Fair Trading, 2001; European Commission, 2008).

Practitioners retain market definition as a first step in abuse of dominance and merger investigations because they view market definition as a preliminary screen: as argued above, market definition is considered a useful tool for assessing market power. But this view ignores that market definition offers a significant other benefit: during market definition, the analyst *identifies and ranks substitutes* for the product sold by the firm(s) under investigation. By treating market definition solely as a means to calculate market share, the analyst foregoes a large chunk of substitution information relevant to an analysis of the competitive effects of either mergers or allegedly anti-competitive practices.

Market definition can be useful under an effects-based approach to abuse of dominance, which requires the analyst to link supposedly anti-competitive behaviour with market effect. Substitution patterns are central to this link: anti-competitive behaviour requires the use of market power, which only exists in the absence of meaningful competitors or the threat of their entry. Market definition therefore assists in assessing the feasibility and possible effects of an allegedly anti-competitive practice – the heart of a competition investigation.

Substitution patterns, and the ranking of substitutes in particular, are also directly useful to a merger investigation. The criticism of market definition as a means of assessing market power usually centres on the fact that market share (i.e. size) is only one factor determining market power. Instead, there is a need for assessing *effective* competitive constraints, i.e. also controlling for countervailing power and the presence or absence of barriers to entry – as these factors can significantly alter the extent of market power, regardless of the size of the firm. But when seen as a ranking of substitutes, rather than an exercise in drawing explicit boundaries, market definition can also contribute in assessing the efficacy of constraints. For example, by ranking substitutes, market definition can help to keep track of the elasticity of the 'fringe supply' – those firms that lie towards the outer boundaries of the market. Fringe supply, in turn, is useful in predicting the likely effects of a market (Blair and Kaserman, 2009: 108).

Market definition, then, continues to be relevant to both merger cases and investigations related to anti-competitive conduct. This is because market definition is not only a first step in measuring market power, but also a first step in organising substitution information central to evaluating competitive effects. While market definition is relevant to modern competition policy, it is necessary to consider its empirical and conceptual challenges. These challenges constitute the research problem of this dissertation. The next section further unpacks the problem by considering the limitations of existing empirical tools for market definition and by identifying two research questions related to quantitative tools for market definition. The subsequent section considers the conceptual challenge of accounting for uncertainty when defining markets, identifying a third research question related to the use of statistical decision theory for market definition.

1.4 Empirical challenges facing quantitative tools for market definition

Quantitative tools for market definition are relatively recent innovations in competition policy. The first significant use of quantitative techniques in market definition in the EU occurred in the 1992 Nestlé/Perrier case (Neven, 2006). In this case, economists used simple price correlation statistics to test whether the relevant market was limited to still water, or should be extended to include all types of bottled water and perhaps even all non-alcoholic drinks (Davis and Garcés, 2010: 171-172). This quantitative technique was introduced for the specific purpose of using an objective criterion for the definition of the relevant market. Of course, as argued later in this dissertation, the price correlation statistic (as any other single quantitative technique) is not necessarily flawless, but served to "inform and improve the decision-making process" (Lexecon, 1994). The "Notice on Market Definition" in the late nineties dramatically accelerated the use of a range of quantitative techniques in the EU, holding that "the systematic identification of the competitive constraints faced by … firms [is] the precise scope of market definition" (Arezzo, 2008: 16).

In the US, quantitative tools have been used relatively longer and a range of quantitative tools have been developed. During the 1980s, quantitative tools for market definition became more sophisticated, mostly due to two developments. Firstly, the rise of cointegration and error-correction models motivated the use of more sophisticated time-series tests of price co-movement. Seminal contributions in this regard include

Horowitz (1981), who suggested the use of long-run equilibrium concepts in a partial adjustment model for market definition, and Slade (1986), who pioneered the use of Granger-causality tests for market definition. Secondly, at about the same time, Scheffman and Spiller (1987) developed the concept of residual demand, which allowed a direct, and arguably more accurate, estimation of price elasticity. Residual demand laid the foundation for the increased use of fully specified IO models for market definition and other purposes.

While both promoted the use of quantitative tools, the two developments contributed to the development of two distinct strands in the literature on market definition tools. The first strand is concerned with extending less data-intensive methods and implicitly promotes using a range of quantitative tools and evidence for market definition. This is an approach consistent with how practitioners approach market definition. The second strand is concerned with finding a single encompassing tool for market definition in the form of a correctly-specified empirical IO model. These IO models can provide a direct estimate of price elasticity that is useful in defining markets, as discussed below, and the models have been applied in a range of competition cases. However, there are challenges to models emanating from this second strand. This dissertation is concerned with the first strand, and it is useful to consider the main practical and theoretical challenges facing IO models in practice.

1.4.1 Challenges facing fully-specified IO models in market definition

As mentioned, practitioners rely on a range of quantitative tools for market definition. The US Merger Guidelines, for example, include the following as evidence for geographic markets: data on shipment patterns, price relationships, transport and distribution costs and excess capacity. Similarly, the Guidelines consider price relationships, product characteristics, and buyer and seller perceptions of product markets (United States Department of Justice and Federal Trade Commission, 2010). More important, in addition to this variety of evidence, price elasticity estimates obtained from empirical IO models are increasingly used. In fact, some practitioners have argued that price elasticity estimates are the *most appropriate* forms of evidence for market definition purposes (see Hosken and Taylor (2004: 465)) and have criticised less data-intensive tools, most notably price tests, for being misleading under a range of conditions. This dissertation investigates the criticisms of price tests and other tools in subsequent chapters. Here it is important to highlight that empirical IO models face important challenges if these are to be used for market definition.

Firstly, competition authorities prefer a range of quantitative and qualitative evidence when defining markets. In fact, in their discussion of market definition tools, Davis and Garcés (2010: 166) note that "qualitative evaluation is universally the starting point of any market definition exercise". In other words, market definition has never been a narrow quantitative exercise aimed at deriving price elasticity estimates from econometric models.

The problem with using elasticity estimates for market definition is that one is implicitly comparing elasticities to some critical threshold. Bishop and Walker (1998: 70) note this tendency and ascribe this to the "precise language in which the test is described", which may be interpreted by econometricians as indicating that markets should be defined by the size of a quantitative estimate of cross-price elasticities alone. The SSNIP test was never intended to be a technical statement on price elasticity, but was intended to describe the importance of evaluating competitive constraints when defining markets. This requires a diverse set of evidence, which can include own- and cross-price elasticity estimates.

Secondly, and perhaps more important to practitioners, there are practical constraints facing competition investigations that favour the use of less sophisticated tools: empirical IO models have significant data, time and capacity requirements that are frequently not met – especially in a developing country context such as South Africa. These constraints are not fully appreciated in the industrial organisation literature, even though competition practitioners have flagged them.

Thirdly, and perhaps less important, the use of price elasticity estimates for market definition purposes face the so-called 'cellophane fallacy' problem in abuse of dominance cases (Forni, 2004). The name 'cellophane fallacy' is derived from the famous US case in which Du Pont, a manufacturer of cellophane, argued, on the basis of a high price elasticity for cellophane, that the material competed with aluminium foil and other packaging in a single market (see Forni (2004: 445-446) and Bishop and Walker (1998: 49)). Typically, the price elasticity of demand is less than unity for lower prices and greater than unity for higher prices. But at which price should elasticity be evaluated for market definition purposes? Usually current market prices are used. However, in a market where firms possess pricing power, the prevailing price will be above the competitive price. Consequently, the corresponding higher elasticity (compared with the competitive situation) will indicate incorrectly that the firm does not have market power. Analysts foreseeing the problem may opt to use a lower price, but such an action leads to circular reasoning. When the purpose of the analysis is to evaluate the possible abuse of market power by a firm, the very goal of defining the market is to ultimately *assess* such market power. Hence, any assumption that the prevailing price is too high indicates that the analyst holds a prior view of market power, before it has been confirmed. Therefore, the use of price elasticity in non-merger competition investigations may be theoretically problematic.

The above discussion indicates that, while empirical IO models can be very useful in defining markets, they are best used as part of a larger investigation based on a range of tools. This dissertation therefore aims to expand the existing quantitative toolkit by adding two empirical tools that are less data-intensive, but nevertheless take account of statistical and econometric advances over the past two decades. Given the specific data and modelling challenges in the contexts of developing countries, this dissertation develops these tools using recent South African competition cases. The following section therefore considers, first, the type of market definition tools adopted historically in South African competition investigations and, second, briefly outlines two research questions aimed at developing two new tools.

1.4.2 Tools for market definition in South Africa

As discussed previously, the 1998 Competition Act established a comprehensive competition policy regime in post-Apartheid South Africa. In the area of merger control in particular, the new Act requires an elaborate and systematic scrutiny of mergers against quite rigorous standards (Roberts, 2004). This stimulated the increased the use of economic analysis and tools in competition practice, and is also reflected in the larger scholarly interest in competition matters (see, for example, the special section on competition policy in the *South African Journal of Economic and Management Sciences* (Roberts, Klaaren and Moodaliyar, 2008)).

Competition policy analysis and tools in South Africa are influenced by approaches and tools in other jurisdictions, but practitioners have been slow to adopt the sophisticated quantitative models increasingly employed in other jurisdictions. Practitioners seem to prefer an eclectic approach that combines a variety of quantitative and qualitative evidence. While modern game-theoretic oligopoly models can be used to directly simulate competitive effects without recourse to market definition, South African practitioners prefer that sophisticated models form part of a broader analysis that also includes the traditional market definition and market power assessment. An example from the literature is Mncube and Ratshisusu (2010), who note that "merger simulation models do not necessarily allow merger analysts to avoid the competitive effects analysis relating to the relevant market, nor do they necessarily provide greater precision to merger control". This emphasis on broad analysis is further supported by the practical problems of capacity, data and time constraints in the South African context.

Practitioners employ a variety of quantitative techniques for market definition in South Africa and Table 1 summarises four papers explicitly dealing with quantitative tools for market definition. Three of the four papers focus on time-series tests of price co-movement with specific application in competition investigations over the past decade.

Table 1: Quantitative tools for market definition studied in the South African literature

Paper	**Quantitative technique**	**Case applied (if applicable)**
Mncube, Khumalo, Mokolo and Nijisane (2008)	Price correlation and univariate stationarity tests on price ratios	Merger between South African steel manufacturers (anonymous)
Boshoff (2007)	Univariate and panel stationarity tests on price ratios	Abuse of dominance complaint against South African dairy processor (anonymous)
Holden (2007)	Armington elasticities for geographic market definition	Range of industries, not specific competition cases
Lexecon (2003)	Stationarity tests on price ratios	Merger between (then) wax producer Schümann Sasol and candle manufacturer Price's Daelite

Given the historical reliance on price tests as quantitative tools for market definition in South Africa, it is useful to situate the two new tools developed in this dissertation within a price context. Firstly, the dissertation develops a new empirical tool for market definition using non-price data. Such a tool can be useful for market definition in markets where systematic price data is difficult to obtain or is unreliable. For example, in many media markets, including markets for newspapers and radio stations, consumers do not pay for products. Under these conditions, it may be useful to consider evidence other than price, specifically consumer characteristics, to investigate the contours of the market. Secondly, the dissertation suggests improved empirical tools when price data is available. As discussed, price tests are frequently criticised, but some of these criticisms can be dealt with by modern econometric procedures and by an approach that views different price tests as complements and not substitutes. The following subsections formally state the research questions underlying these two empirical tools.

1.4.3 Research question 1: Consumer characteristics and cluster analysis for market definition

Chapter 2 considers the first research question of the dissertation: can consumer characteristics assist in market definition and what tools can be used to compare the consumer characteristics of different products?

Consumer characteristics refer to the demographic and income profiles of consumers and can be useful in defining markets for differentiated products, as these characteristics influence the extent to which the consumer is willing to switch to alternative products. The empirical IO literature has long recognised the utility of consumer characteristics in studying substitution, and data on consumer characteristics is often incorporated into empirical IO models to improve elasticity estimates (seminal contributions in this regard include Berry, Levinsohn and Pakes (1995; 2004)). Unfortunately, consumer characteristics are not always readily available at the individual level. Despite this limitation, even an average profile of consumers of different products can be useful in determining substitutability among these products (see Petrin (2002) for a prominent application). This suggests that consumer characteristics can assist in market definition.

From a statistical perspective, when consumer characteristics of products are similar, the vectors of consumer characteristics for the products will tend to cluster together. Therefore, a useful tool for comparing consumer characteristics is statistical cluster analysis. Chapter 2 explores the various clustering techniques offered in the statistical learning literature and demonstrates their application to market definition in a radio merger case recently concluded in South Africa.

1.4.4 Research question 2: Price-time-series tests for market definition

Chapter 3 explores a second research question: are the various price-time-series tests for market definition consistent and can they be improved to reflect recent time-series developments?

Although used fairly often, price tests have been heavily criticised (see Coe and Krause (2008) for a recent simulation-based criticism). Chapter 3 considers the conceptual and empirical criticisms, but suggests that recent small sample improvements to time-series tests (including better unit-root and cointegration tests), and a strategy that is sensitive to the specific and complementary hypotheses underlying different price tests, may yet render price tests useful in defining markets. Specifically, Chapter 3 uses milk price data from a recent competition investigation to demonstrate how correlation analysis, Granger-causality tests, unit-root tests, and bounds tests using autoregressive-distributed-lag models, can assist in defining geographic markets.

1.5 The conceptual challenge of uncertainty in market definition

The above discussion suggests that empirical tools for market definition face challenges that inevitably create uncertainty about market definition conclusions. The binary approach commonly adopted when defining markets is an important source of uncertainty: competition law requires a market with clear boundaries, and practitioners therefore tend to classify a product as either ‘in’ or ‘out’ of the market.

Binary classification is particularly problematic in differentiated product markets, which explains why market definition is frequently contentious. In response, this dissertation develops a statistical decision rule for market definition, which explicates the relative roles and interplay of the underlying preferences of the decision maker, case evidence and prior probabilities (based on past cases and literature). The following subsection formally states the research question underlying this conceptual tool.

1.5.1 Research question 3: Market definition as a statistical decision rule

If market definition is seen as a binary classification problem, it will inevitably face significant misclassification probabilities (for incorrect inclusion and incorrect exclusion). Chapter 4 therefore considers a third research question: given that market definition occurs under uncertainty, can one use a statistical decision rule to unpack the various factors important to a market definition decision?

Chapter 4 suggests a Bayesian decision rule for including or excluding a product from the relevant market. The Bayesian framework includes prior beliefs and loss functions to deal with uncertainty in market definition. Prior beliefs are included in the form of the weight of substitutability evidence from extant literature and previous cases, while the loss function explicitly states the costs assigned to overly narrow or overly broad market definitions. Prior probabilities and losses are then weighed with substitutability evidence from the case to determine the market boundaries.

It is limiting to interpret the Bayesian framework as best implemented using a single econometric model. The Bayesian rule favours the use of a variety of tools and evidence for rational market definition. Chapter 4 uses the flexible product market definition in a recent South African radio merger case to illustrate the potential benefits of a decision rule approach to market definition.

1.6 Summary

Economists have played an important role in shaping the conceptual framework of market definition in different competition policy jurisdictions. At the same time, economists have worked on improving the empirical analysis of substitutability. This dissertation follows these trends and develops new conceptual and empirical tools to enhance market definition in competition law investigations. Specifically, as discussed above, the following three chapters each consider a research question:

(i) In Chapter 2, research question 1: Can consumer characteristics assist in market definition and what tools can be used to compare the consumer characteristics of different products?

(ii) In Chapter 3, research question 2: Are the various price-time-series tests for market definition consistent and can they be improved to reflect recent time-series developments?

(iii) In Chapter 4, research question 3: Given that market definition occurs under uncertainty, can one use a statistical decision rule to unpack the various factors important to a market definition decision?

Chapter 2 therefore develops a new market definition tool based on consumer characteristics and cluster analysis, while Chapter 3 introduces new price-time-series tests and argues for the use of a range of price tests. The conceptual contribution is contained in Chapter 4, which uses statistical decision theory to derive a Bayesian decision rule that can assist practitioners in defining markets, by making explicit the roles of case evidence, prior probabilities and the loss functions of decision makers. Chapter 5 concludes by summarising the main arguments of Chapters 2 to 4 and identifying future research opportunities.

Chapter 2

Market Definition Using Statistical Learning Techniques and Consumer Characteristics

Market definition relies on empirical evidence of substitution patterns. Cross-price elasticity estimates are preferred forms of substitutability evidence, due to advances in econometric modelling in the field of industrial organisation (IO) over the past two decades. However, as mentioned in Chapter 1, the 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, combining different tools significantly expands the available information set and broadens the basis for inferences.

This chapter proposes a complement to the existing set of tools, namely an analysis of consumer characteristics for market definition purposes. Specifically, the chapter considers the following research question: can consumer characteristics assist in market definition and what tools can be used to compare the consumer characteristics of different products?

To this end, the chapter first presents a consideration of the extent to which a study of consumer characteristics is consistent with the concept of a market in competition law investigations. Thereafter, the chapter shows how cluster analysis (a type of statistical learning technique) can be used to identify meaningful groups of substitutes based on the homogeneity of their consumer characteristics. Cluster analysis enforces consistency in the analysis of characteristics, while recent advances now also allow statistical inference to ensure robust conclusions. The relatively easy implementation and graphical output, which assist interpretation, further recommend cluster analysis as a tool for market definition. Finally, cluster analysis on consumer characteristics data is applied to the market definition problem in a recent radio merger case in South Africa.

2.1 Can consumer characteristics assist in defining markets?

When defining a market, 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 SSNIP thought experiment. Chapter 1 argues that the SSNIP test focuses on the consumer, i.e. demand-side behaviour. This consumer focus suggests that empirical tools analysing consumer-related data can be particularly useful in market definition, and the competition literature has developed a number of these tools: consumer surveys providing diversion ratios (Katz and Shapiro, 2003; O'Brien and Wickelgren, 2003; Farrell and Shapiro, 2010) and econometric IO models providing price elasticity estimates (Froeb and Werden, 1991; Ivaldi and Verboven, 2005) are useful tools as they are directly relevant to the SSNIP test. However, especially the econometric IO models are not always feasible to use. Furthermore, even if these models were feasible, Chapter 1 details theoretical and practical challenges justifying the use of a range of complementary tools for market definition purposes.

Suggested in this chapter is a complementary empirical tool for market definition, namely the study of heterogeneity in consumer characteristics. The rationale for such a market definition tool comes from extant competition policy literature, which suggests that it is useful to consider the demographic profile of consumers of differentiated products when studying substitutability. Coate and Fischer (2008: 33) note, for example, how the substitutability of differentiated products depends on the class of consumer: “two products that appear relatively similar to one customer may be highly differentiated to another”. In fact, consumer characteristics have played an important role in improving elasticity estimates in modern 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 the 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); 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 data on individual consumer characteristics may be useful for market definition purposes (Berry et al., 2004), data constraints may limit their use in competition investigations. However, while consumer-level data significantly improves estimates because of its 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, a study of aggregate consumer characteristics can be a complementary tool for market definition.

A study of consumer characteristics can be particularly useful for market definition in two types of competition investigation. Firstly, heterogeneity of 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 to define markets using data on consumer characteristics. This data is frequently available from marketing research firms (in the form of consumer surveys or market studies) and/or from rich sales data at the firm itself (companies increasingly collect and store consumer data).

Secondly, heterogeneity of consumer characteristics can help to define markets 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.

Even in other types of competition investigations, a study of the heterogeneity of consumer characteristics remains a useful additional tool capable of providing 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 the focus is on a media market where consumers do not pay in monetary terms and where the product is targeted at a particular consumer group. Before considering this application, the following section discusses the tool called cluster analysis, which is used to statistically analyse data on consumer characteristics.

2.2 Cluster analysis

The study of consumer characteristics in defining markets proceeds from the assumption that the differences in the consumer characteristics of two products shed light on the cross-price elasticity of demand for the two products. Suppose two firms selling products a and b propose to merge. To define the relevant market(s) for this merger investigation, 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 $\boldsymbol{c}_a$ and $\boldsymbol{c}_b$. The similarity of the consumer characteristics for the two products is then captured by the distance between $\boldsymbol{c}_a$ and $\boldsymbol{c}_b$, called $d(\boldsymbol{c}_a, \boldsymbol{c}_b)$. This chapter argues that $d(\boldsymbol{c}_a, \boldsymbol{c}_b)$ approximates the cross-price elasticity between products a and b.

When $d(\boldsymbol{c}_a, \boldsymbol{c}_b)$ is small, the cross-price elasticity between the two products is high, and vice versa. 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 formally identify such clusters. One set of tools frequently used to group multivariate data is cluster analysis (Lorr, 1983; Kaufman and Rousseeuw, 1990). This chapter focuses on hierarchical cluster analysis as the statistical tool for studying multivariate data on consumer characteristics.

Before considering the technique itself, it is important to note up-front that cluster analysis identifies group structures in data on the basis of criteria specified by the analyst. 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). The grouping structures identified in the data are therefore never 'objective' or 'natural', but the result of a particular similarity concept employed. 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). Cluster analysis is therefore not "a graphical summary of the data itself [but] a description of the results of the algorithms" (Hastie, Tibshirani and Friedman, 2001: 475). The following subsections therefore discuss the various choices available for the distance metric and linkage methods, before proceeding to a discussion of the process of hierarchical cluster analysis.

2.2.1 Distance

The choice of distance metric is determined by the analyst's tolerance for dissimilarity. In defining markets, the analyst wants to group product A (or, if it is a merger between two firms, products A and B) with competitors of the product(s) that satisfy the SSNIP test. The analyst will seek to avoid defining the market as overly narrow or overly broad (see Chapter 5 for a discussion of loss functions). Therefore, a distance metric that does not overstate differences among substitutes may be preferable.

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 larger differences than to smaller ones, so variables with larger orders of variation will dominate more stable variables in the distance metric:

$$d(\boldsymbol{c}_a, \boldsymbol{c}_b) = [(\boldsymbol{c}_a - \boldsymbol{c}_b)'(\boldsymbol{c}_a - \boldsymbol{c}_b)]^{\frac{1}{2}}$$

where $\boldsymbol{c}_a$ and $\boldsymbol{c}_b$ have been mean-centred.

The emphasis on larger orders of variation is useful when defining the market, as the competition analyst may want to include products in the market that are slightly different, but may be more wary of including them when they differ markedly. One could replace the squared distance with an absolute deviation distance, but this would reduce the emphasis on greater differences, and hence, this chapter 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:

$$d(\boldsymbol{c}_a, \boldsymbol{c}_b) = [(\boldsymbol{c}_a - \boldsymbol{c}_b)'\boldsymbol{\Lambda}^{-\boldsymbol{1}}(\boldsymbol{c}_a - \boldsymbol{c}_b)]^{\frac{1}{2}}$$

where $\boldsymbol{\Lambda}$ 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 – it weighs the different variables with their respective sample variances and transforms the system of correlated coordinates into a system of uncorrelated coordinates:

$$d(\boldsymbol{c}_a, \boldsymbol{c}_b) = [(\boldsymbol{c}_a - \boldsymbol{c}_b)'\boldsymbol{S}^{-\boldsymbol{1}}(\boldsymbol{c}_a - \boldsymbol{c}_b)]^{\frac{1}{2}}$$

where $\boldsymbol{S}$ is the sample covariance matrix.

This chapter uses statistical distance in the empirical application. Note further that cluster analysis requires actual distances, calculated from underlying data. In many competition cases, consumer surveys are conducted where participants are asked to judge the degree of substitutability among different products. These subjective "distance" judgments are not directly useful for cluster analysis as they frequently violate triangle inequality[4] (Hastie et al., 2001: 455). Cluster analysis applied to subjective distances would therefore be misleading, and the data must be transformed into dissimilarities for cluster analysis purposes. The transformed distance matrix obtained from this data may also be nonsymmetric and will have to be transformed into a symmetric form.

2.2.2 Linkage method

The distance metric is a measure of the similarity of observations. Cluster analysis also requires a measure of the 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 merging into a single cluster. Cluster linkage methods differ according to the size bias: some linkage methods have a higher tolerance level for dissimilarity within merged clusters and will tend to suggest larger clusters.

It is advisable to 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 each successively larger cluster groups together less similar items. Therefore, the overall variance within clusters should increase as one moves up the tree. Inversion refers to the violation of this property, where a cluster at a particular height has a *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 producing inversion for particular types of data and are best avoided. The problem of inversion raises the general question of which linkage method is to be preferred. 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

[4] Triangle inequality requires that $d(\boldsymbol{c}_a, \boldsymbol{c}_b) \leq d\left(\boldsymbol{c}_a, \boldsymbol{c}_j\right) + d\left(\boldsymbol{c}_b, \boldsymbol{c}_j\right)$ for all $j = 1, \ldots, N$.

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 merging of two clusters requires only that one of the observations in each of the clusters is sufficiently close to the other. 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 "closeness" are to the analyst (Johnson, 1967; Lorr, 1983; Hastie 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. From a statistical inference perspective, the closeness criterion therefore aims to minimize Type I error (incorrectly excluding substitutes), while the compactness criterion aims to minimize Type II error (incorrectly including substitutes). Chapter 4 further explores these errors and their associated losses in market definition.

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 et al., 2001: 477): even if a strictly increasing transformation is applied so that the relative ordering of distances is preserved, the answers may change for different distance metrics. The single and complete linkage methods require only an ordering of the distances and are invariant to strictly increasing transformations (Hastie 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 an attempt at identifying the local modes in a joint probability distribution, and Kelly and Rice (1990) show that only clusters produced by the average linkage method bear any relation to the characteristics of a 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 further on 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 the market.

2.2.3 Hierarchical cluster analysis and its relation to market definition

The linkage method and distance metric are the similarity measures that dictate the type of group structure suggested by 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 a 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 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 et al., 2001). This method generates the actual clusters as its output.

Hierarchical cluster analysis does not generate clusters, but produces a hierarchical tree diagram, called a dendrogram (see Figure 3, Figure 5 and Figure 7 in the empirical application). 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 terminal nodes have a 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 based on one of the distance concepts discussed earlier.

(iii) Search for the most similar pair of nodes U and V (i.e. the smallest distance in the distance matrix), and merge these nodes to form a new node UV.

(iv) Calculate a new distance matrix for the reduced set of nodes, using the particular linkage method to calculate distances among original nodes and the newly formed node. The distance between node UV and another node W, $\boldsymbol{d}_{(UV)W}$ is calculated as follows:

Average linkage: $\boldsymbol{d}_{(UV)W} = \frac{\sum_i \sum_j \boldsymbol{d}_{ij}}{N_{(UV)} N_W}$ where $N_{(UV)}$ is the number of elements in node UV and N_W is the number of elements in node W

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

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

(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 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, at minimum should be included. 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 clusters (Tibshirani, Walther and Hastie, 2001). The following section explores optimality criteria and techniques for identifying the optimal number of markets.

2.2.4 Forming clusters from the dendrogram

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 an 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 analyses for different datasets generate 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 a lower variance within clusters compared with 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 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:

$$D_r = \sum_{a,b \in C_r} d(\boldsymbol{c}_a, \boldsymbol{c}_b)$$

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

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

W_k is generally *decreasing* 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 the 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 et al., 2001).

The graphical approach to identifying 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:

$$gap(k) = E^*(\log(W_k)) - \log(W_k)$$

The gap statistic outperforms alternative methods for finding the optimal number of clusters and performs well in cases where the data is not naturally separated into more than one group (Tibshirani et al., 2001). Despite these advantages, the gap statistic does not monotonically decrease 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 is demonstrated in the application.

This chapter implements a 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.

2.2.5 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 a grouping methodology as described above, the practitioner also seeks assurance that the sampling error in the data has not compromised the dendrogram, in whole or in part. The search for robust, generalisable conclusions in competition policy is not new, as the growth of econometrics in empirical industrial organisation 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 development 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: because 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 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 and Holmes, 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 of $O(n^{-1})$ (Shimodaira, 2002).

The literature on bootstrapping for hierarchical clustering developed a range of alternative bootstrap techniques, including multi-scale bootstrapping, to reduce the 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 a second-order size bias of $O(n^{-\frac{1}{2}})$. 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 $n^{-1/2} = 0.22$ 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 bias of $n^{-3/2} = 0.01$.

The preceding discussion has covered the following: the various technical aspects of conducting a hierarchical cluster analysis for market definition purposes, the appropriate choice of distance measure and linkage method, the process of generating the dendrogram and identifying clusters, and bootstrapping for statistical inference. The following two sections demonstrate the tool using data from a recently concluded South African radio merger case: the first section briefly discusses the case and the consumer data, while the second section considers the cluster analysis results.

2.3 **Illustrative case: Primedia / Kaya FM partial radio station merger**

In 2006 the South African media group Primedia together with the 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 (also refer to the discussion in Chapter 4).

The following subsections consider the validity of the case for demonstrating the use of cluster analysis and consumer characteristics in market definition, then the case background, and thereafter the data sources and properties.

2.3.1 Validity of the case

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, listeners are looking for access to this market place based on their specific preferences for news products, music products, talk shows, etc. and they pays with the opportunity cost of their 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. Advertisers are looking for access to the market place so that they may sell information products 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 these 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 profiles 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[5].

Of course, it is always useful to obtain direct evidence on both sides, as an 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 the argument above, consumer characteristics shed light on (i) and (iv), and should ideally be complemented by further information on (ii) and (iii).

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

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, consumers still pay with the opportunity cost of their 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 an alternative measurement of the "notional" price in order to infer cross-price elasticities, which may be difficult to obtain (Europe Economics, 2002). One may argue that price data on the advertiser side is still available. Such price data is not directly relevant to the analysis of the listener-side, but even if it were relevant, advertiser price data was problematic in the Kaya case. While advertiser rate cards (list prices for radio advertisements) were available, closer analysis during the case proceedings revealed that the rate cards bore little relation to the actual prices charged by radio stations. This was also recognised by the Tribunal in their decision: "[T]here is no relationship between rate cards and what the radio stations charge in reality. The rate cards are not reflective of prices" (Competition Tribunal, 2008: 18). Given these constraints, the Primedia/Kaya radio merger is a useful case to demonstrate how an alternative tool in the form of cluster analysis can assist in market delineation.

2.3.2 Background on the product and geographic market

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[6] radio stations, Highveld Stereo and Radio Jacaranda). In early 1997 eight new commercial radio licences were granted for broadcasting in South Africa's three largest 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 competing directly with mainstream privately owned stations and SABC public stations focusing on specific cultural or language groups.

[6] 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.

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

- SABC commercial: 5fm, Metro
- 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

This list excludes so-called "community" radio stations run on a non-profit basis. Since the late 1990s, broadcasting authorities have issued more than 100 community radio licenses across the country. 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 listener numbers are generally very small. These stations do not fundamentally alter the market definition conclusions obtained from the statistical analysis and have been omitted from the cluster analysis. The analysis also excludes Internet radio stations, which had negligible listener numbers at the time.

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). At least two of the parties in the case defined 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. For Gauteng radio listeners who prefer local content, radio stations with a national footprint, and local stations without a Gauteng footprint have to be excluded ('footprint' refers to the proportion of Gauteng province reached by the station signal). However, where general content is preferred, national radio stations with a Gauteng footprint are also plausible competitors for listeners. This chapter focuses on the broader geographic market, which includes local radio stations as well as national radio stations with a Gauteng footprint. The Competition Commission 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 in the developed world, based on US and UK studies (Silk, Klein and Berndt, 2002; Waldfogel, 2002). In contrast, authors working on the advertising side of radio markets find radio advertising to constitute a separate market (Ekelund, Ford and Jackson, 1999). The position in the competition literature is also that radio markets are to be distinguished from other media markets, by virtue of their particularly personal and intrusive nature. This chapter focuses on radio as a distinct product market, following the approach of the competition authorities (Competition Tribunal, 2008: 10). The issue receives further attention in the discussion in Chapter 4 of the role of prior probabilities in market definition.

2.3.3 Data

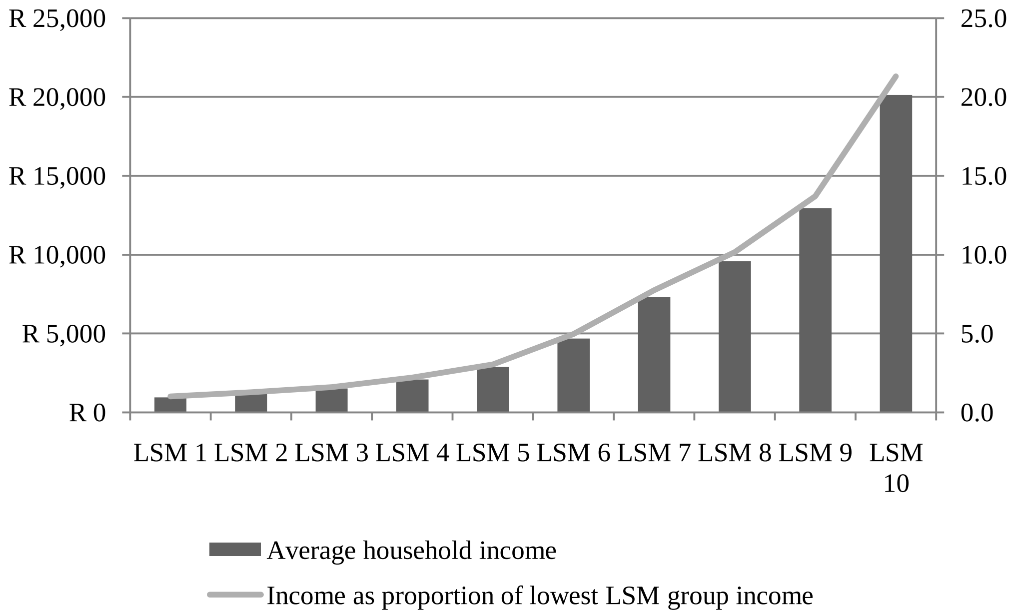
The study relies on a dataset of radio listener characteristics obtained from the South African Advertising Research Foundation's Radio Audience Measurement Survey (RAMS) for 2004/5, the year in which the merger application was first heard before South African competition authorities (South African Advertising Research Foundation, 2005). RAMS is an extensive media research database 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). 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.

RAMS is useful for this chapter, as it contains data on the demographic and income profiles of listeners of each radio station. A survey respondent is defined as a listener of a particular 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. Such a disaggregated analysis is not attempted in this study and the focus is on the overall number of listeners over seven days. This issue did not feature prominently in the merger proceedings.

RAMS does 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 discussed below, they differed on whether to consider this group as a whole for market definition purposes.

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



RAMS contains the number of listeners of each radio station falling into each of the following categories:

Table 2: RAMS data on listener profile of Gauteng radio stations

Listener characteristic	**Description**
Income group	Ten categories: LSM 1 to 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

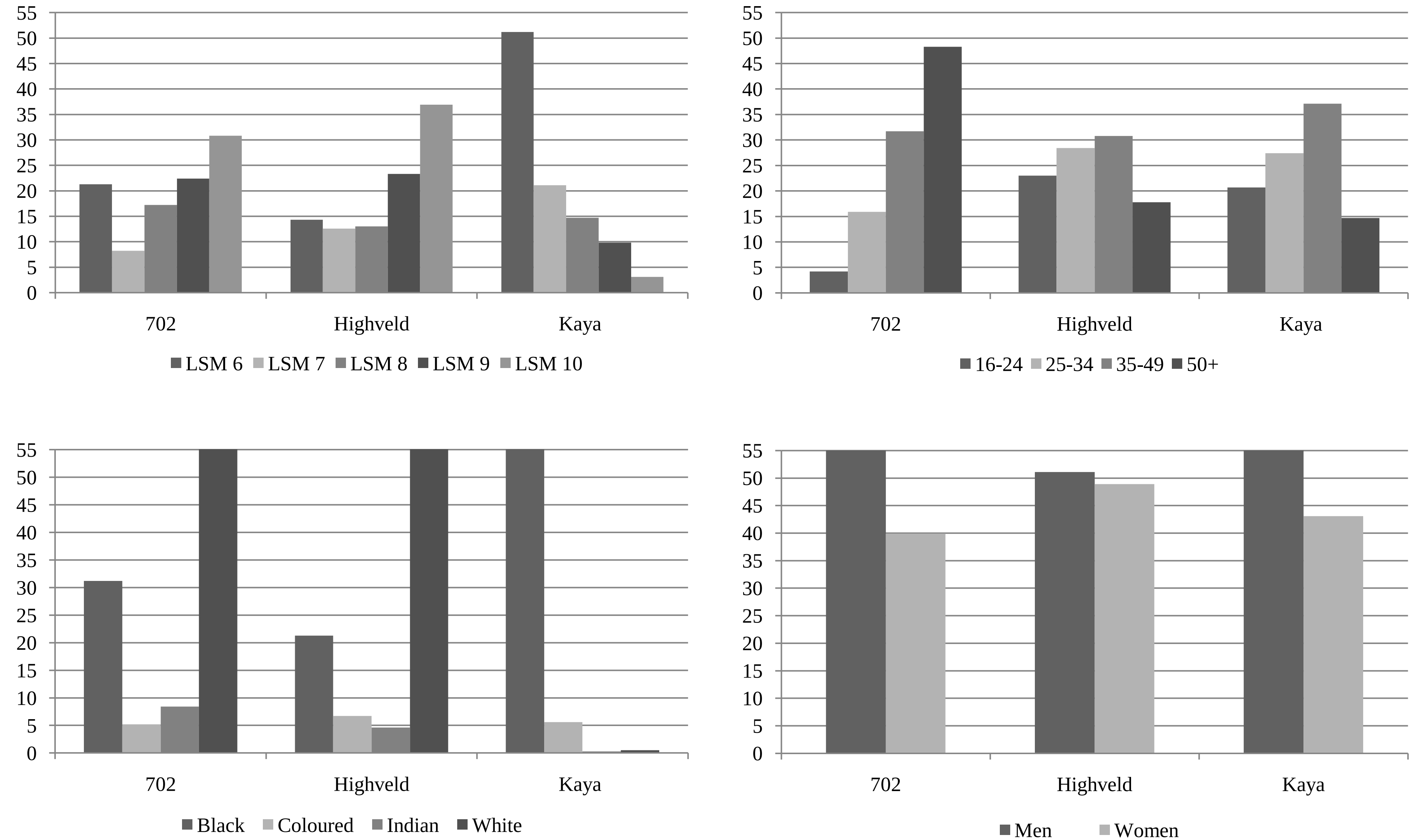
RAMS captures the demographic and income characteristics in *absolute* numbers (thousands of listeners). However, consumer characteristics are better measured in proportions. Absolute numbers would bias a market definition based on cluster analysis, as radio stations with large audiences would be found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

It is important that the observed consumer characteristics represent the true consumer characteristics that matter for the SSNIP, as this ensures that the empirical analysis of consumer characteristics is relevant for market definition purposes. Suppose one considers race and age as the only dimensions in terms of which to compare consumers and, therefore, uncovers significant differences between old White and young Black listeners; yet, if income and sex, rather than race and age, are the more relevant drivers of demand, such a comparison is misleading. 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 (Berry et al., 2004). This difficulty may lead to contention in competition cases surrounding the choice of particular variables.

In the radio case, 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 that can be assumed to be collecting data relevant to its members for targeting purposes, and this confirms the salience of these variables in a demand analysis. Secondly, the 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 includes 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 Kaya 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 included LSM 6 to LSM 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 their limited set of characteristics, which restricts 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.

Figure 2: Listener profiles per LSM (income), age, race and sex for merging radio stations



2.3.4 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 income and racial differences, especially between Highveld and Kaya listeners. However, while the stations differ significantly on income and racial profiles, their listeners are more similar in terms of age – especially when compared to 702. This illustrates the problem of integrating univariate conclusions: a set of diverging univariate conclusions on income, race, age, etc. must be integrated 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 chapter suggests cluster analysis, a systematic grouping method that can also deal with multivariate observations.

2.4 Results of cluster analysis

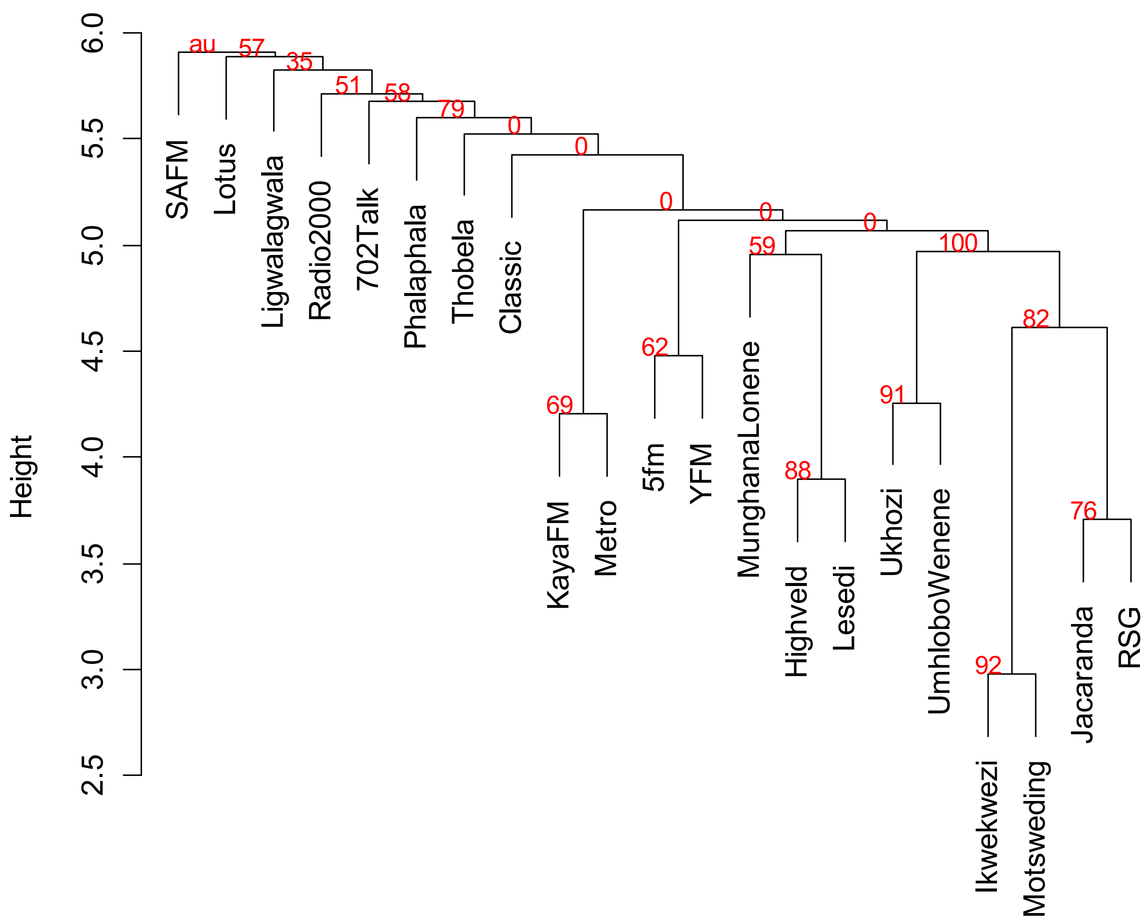
This section shows the results of applying a hierarchical cluster analysis to the data on radio listener characteristics described in the previous section. Results are reported separately for the average, single and complete linkage methods and focus on the relationship between Highveld and Kaya rather than on 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.

2.4.1 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 approximately unbiased bootstrap probabilities. The dendrogram reveals important evidence for market definition purposes. Some radio stations are difficult to group and are 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. More important, the dendrogram 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 and RSG, then Highveld and Lesedi, and then Kaya and Metro. Importantly, Kaya and Highveld only cluster together at a later stage, suggesting that a number of competitors separate these two radio stations. This feature assists with answering one of the important market definition issues in the case, namely which stations to include, *at minimum*, with Highveld and

Kaya if one assumes that these two stations share the same market. Figure 3 suggests that, if Highveld and Kaya share a market, a number of other stations have to be included: Metro, 5FM, YFM, Munghana Lonene, Lesedi, Ukhozi, Umhlobo Wenene, Ikwekwezi, Motsweding, Jacaranda and RSG.

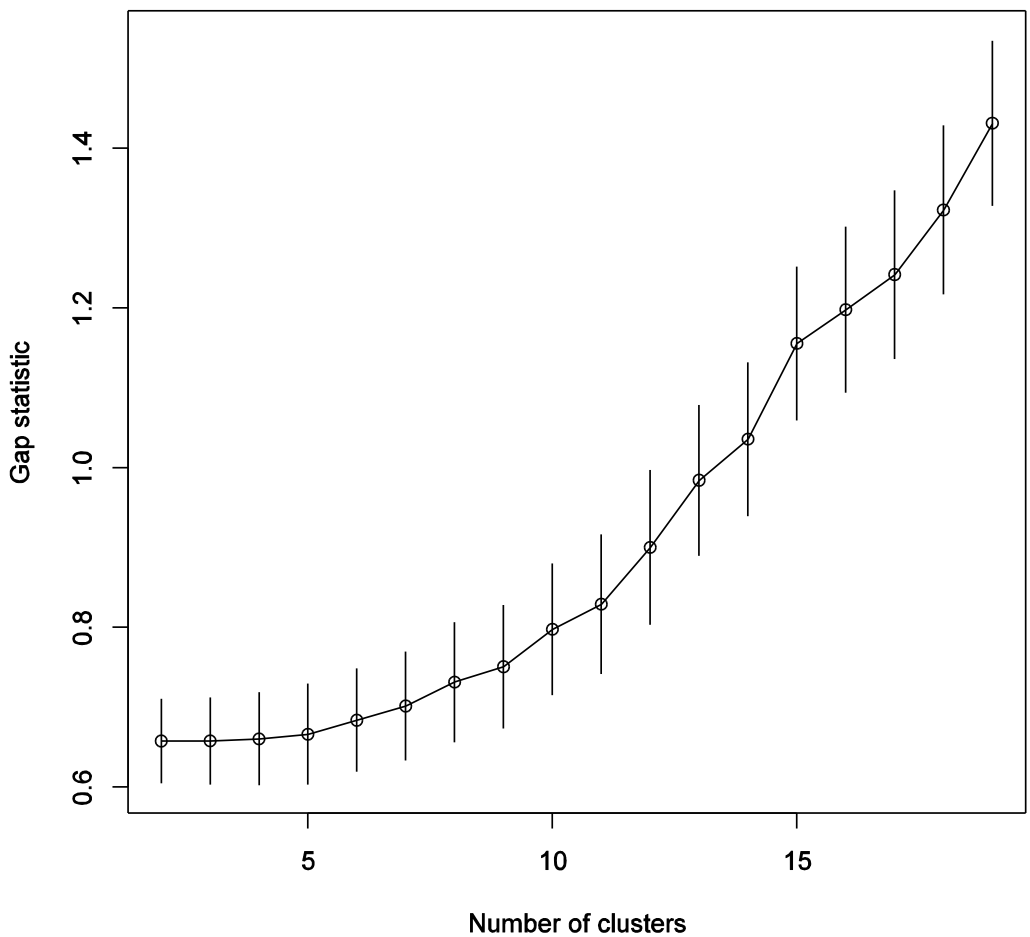
Figure 3: Dendrogram of radio listener characteristics, based on average linkage method and statistical distance



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 in the event of small changes in the data. Instead, it is better to identify those parts of the dendrogram that are more robust. It is clear that Highveld shares a robust cluster with a set of other competitors – indicating that it is important to first 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 was generated using a Monte Carlo simulation of 1000 repetitions. Figure 4 plots the gap statistic for each assumed number of clusters.

Figure 4: Gap statistic for dendrogram of radio listener characteristics, based on average linkage method and statistical distance



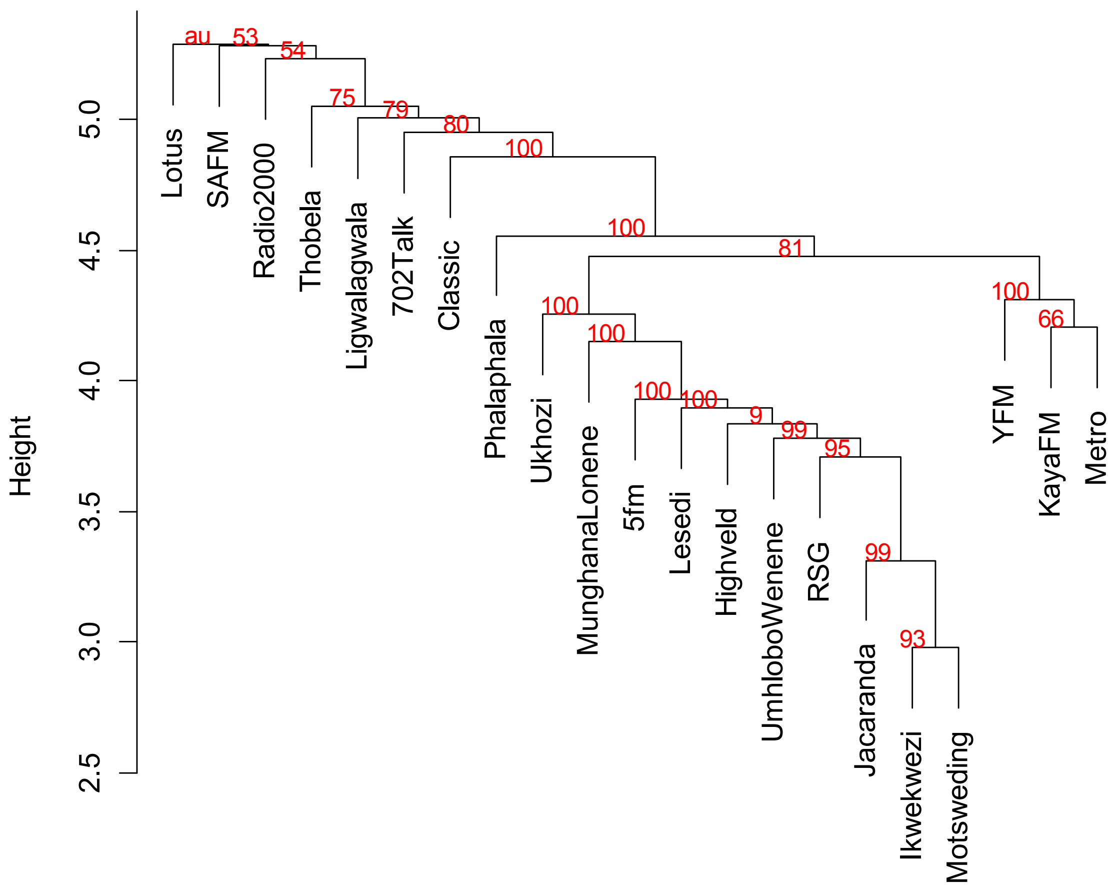
The gap statistic suggests that the data is not easily separable into clearly identifiable clusters: the gap statistic continues to rise as the number of clusters increase. Nevertheless, the gap statistic does indicate a maximum at two clusters, as it does not increase significantly when adding a third, fourth or fifth cluster. 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 et al., 2001). Therefore, if analysts are interested in identifying a large number of tight clusters, they 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.

2.4.2 Single linkage method

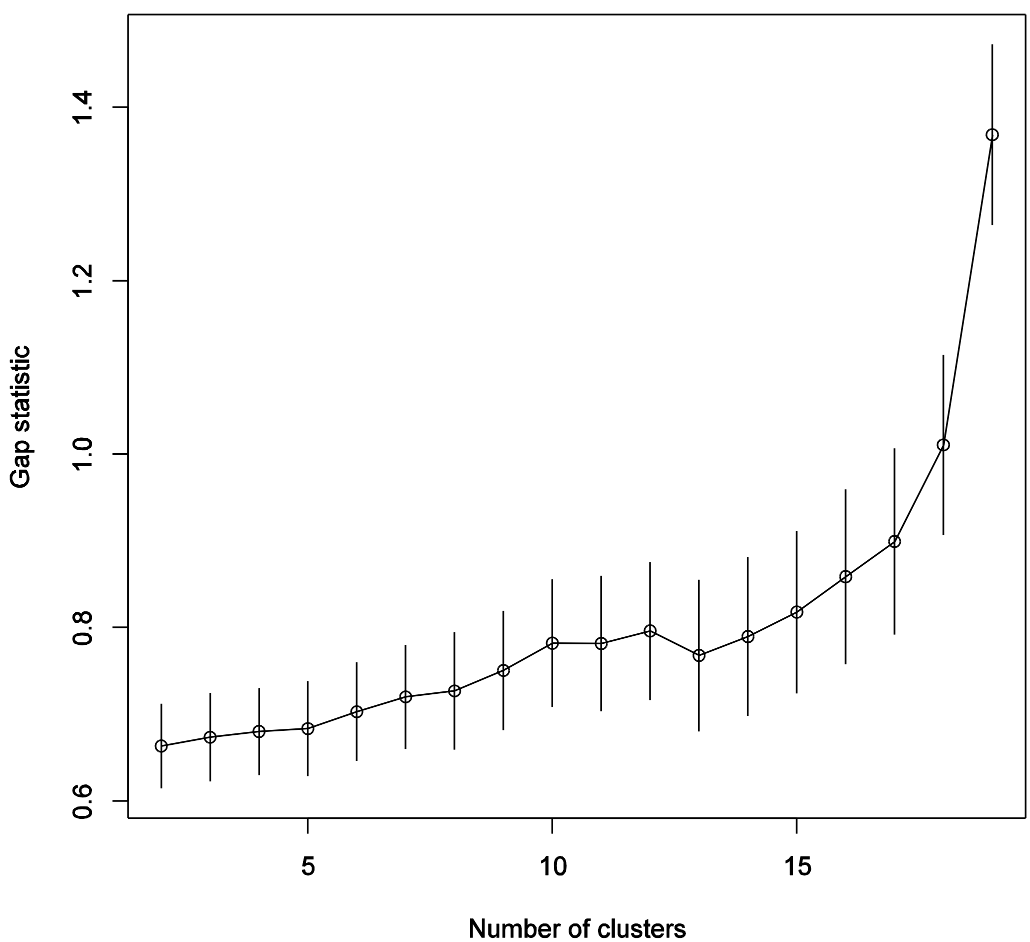
Figure 5 repeats the dendrogram using the single linkage method:

Figure 5: Dendrogram of radio listener characteristics, based on single linkage method and statistical distance



The results for the single and average linkage methods are similar. Figure 5 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 6.

Figure 6: Gap statistic for dendrogram of radio listener characteristics, based on single linkage method and statistical distance

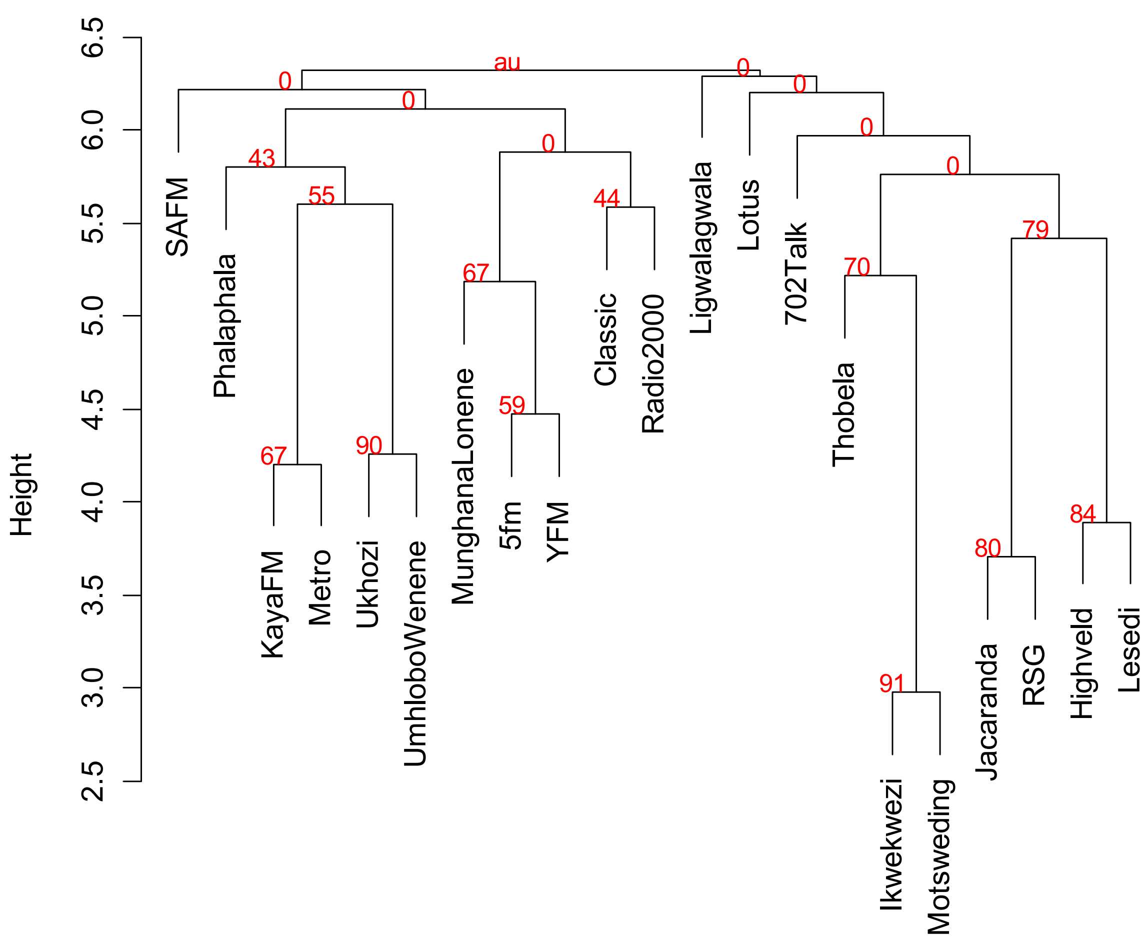


Similar to the average linkage method, the gap statistic also seems to grow quite slowly for clusters 2 through 5 – but does not reach a local maximum. Figure 6 suggests a local maximum at 13 clusters: the gap statistic rises significantly from 2 through 12, but declines for 13. Moving beyond 13, the 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 to attempt to group radio stations into a specific number of clusters (see Chapter 4 for a discussion of the Tribunal's market definition, where they acknowledge this problem).

2.4.3 Complete linkage method

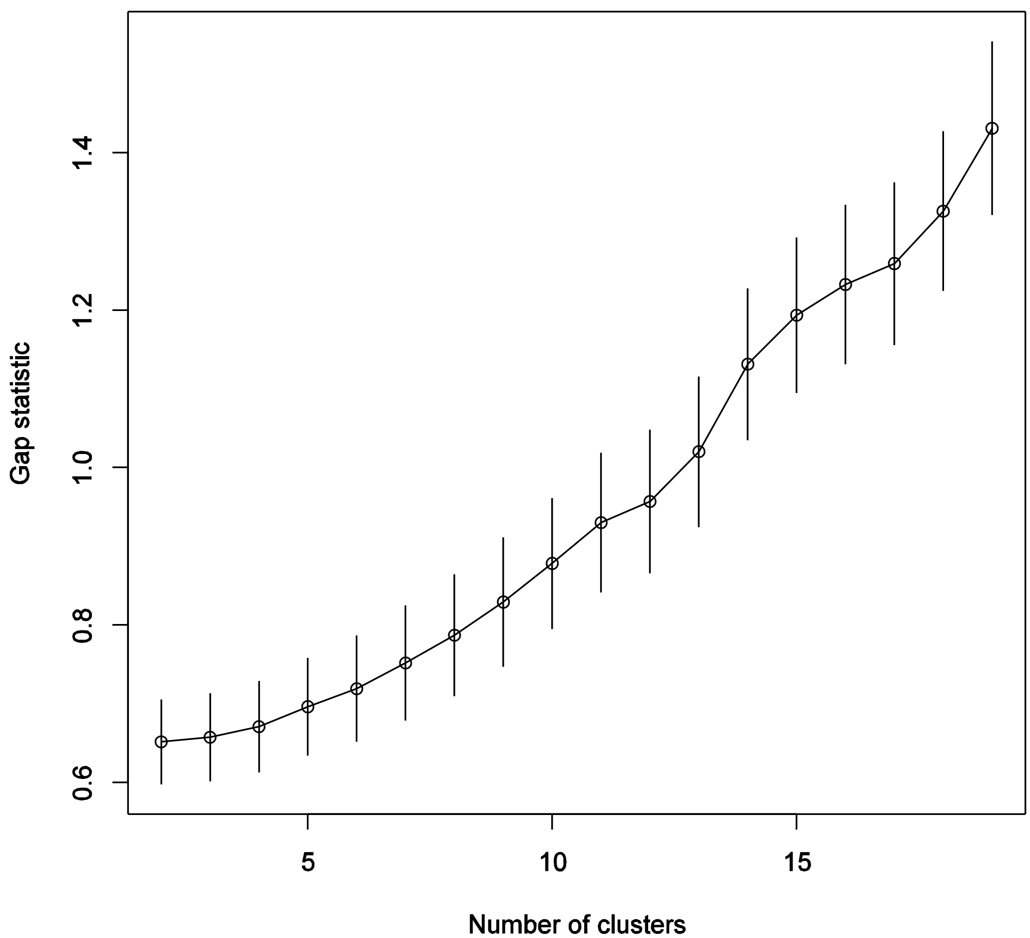
Figure 7 presents the dendrogram for the complete linkage method, including bootstrap probabilities:

Figure 7: Dendrogram of radio listener characteristics, based on complete linkage method and statistical distance



This dendrogram is quite different from that of either 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 the clusters were 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. 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 8.

Figure 8: Gap statistic for dendrogram of radio listener characteristics, based on complete linkage method and statistical distance



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 support this and does not reach any local maximum, although the statistic grows slowly from two to three clusters.

2.4.4 The gap statistic and uncertainty

This chapter 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 $gap(k) \geq gap(k+1) - s(k+1)$ where $s(k) = sd(k)/\sqrt{1+\frac{1}{B}}$ 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 k. Figures B1, B2 and B3 in Appendix B 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 confirms the earlier point that the data for this case is better investigated asking specific substitution questions rather than attempting to delineate specific markets.

2.5 Summary of cluster results and corroborating evidence

The dendrograms presented above are useful graphical tools that reveal a significant amount of information on substitutability. This section summarises the conclusions from each linkage method. Table 3 shows which other stations would populate a market that included both Highveld and Kaya. The table also allows an assessment of how robust the suggested markets are, if one were to limit conclusions to the *largest* possible clusters with an 80% or higher bootstrap confidence level.

Table 3 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 more than an 80% probability (in a bootstrap sense) that the entire cluster would 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 with small data changes Kaya would 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 would 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 appear to organise into robust clusters under the complete method).

Table 3: 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

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. However, the markets defined here are based on a systematic analysis of data, while the Tribunal did not attempt any formal quantitative evaluation. Cluster analysis could therefore have assisted the Tribunal in ranking substitutes and drawing market boundaries.

The cluster analysis produces the same conclusions on market definition using the average and single linkage methods (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 potentially 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. Chapter 4 deals explicitly with the roles of uncertainty and diversity of evidence in defining markets.

2.5.1 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 the radio market. Both sides were studied during the merger hearings, but in this chapter the interest lies with the extent of the listener side of the market in particular. However, other data from RAMS offers potentially useful corroborating evidence. The data suggests that, on average, South African radio listeners tune in to two to three radio stations per day. In fact, RAMS enables 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 is shown in Table 4.

Table 4 confirms the findings from the cluster analysis. Firstly, the cluster analysis and switching analysis suggest virtually the same substitutes for Highveld and Kaya. Secondly, the switching analysis 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.

Table 4: 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

The cluster analysis suggests that Highveld and Kaya may be in the same overall market, but the dendrogram shows 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 regulated, and radio station licences are granted by the Independent Communications Authority of SA (ICASA). The radio licence agreement stipulates the format of a radio station, usually specifying the type of content and/or language to be used. In its merger decision, the Tribunal summarises the required radio formats for Highveld, 702 and Kaya as follows:

Table 5: Radio formats based on licence conditions

Radio station	Format
Highveld	"A commercial adult contemporary radio station"
702	"A commercial talk radio station"
Kaya	"An African-focused adult contemporary / jazz format with a 60% music and 40% talk format"

Source: Competition Tribunal (2008: 4)

Although both Highveld and Kaya hold licences to play "adult contemporary" music, the Kaya licence constrains the format to focus 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.

2.6 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[7] 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 chapter, 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 to develop a set of

[7] 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.

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 stated 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 chapter 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 are 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 greater 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 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 further investigate 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 the 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. It is argued here that the cluster analysis still allows one to assess where marginal consumers may turn to, as it is based on identifying substitutes for the product under investigation on the basis of the similarity of products in terms of some rather than all consumer characteristics. Nevertheless, the average profile may still result in an underestimation of the extent of substitution between some radio stations if the consumer most sensitive to price is quite different from the average consumer.

Finally, cluster analysis does not recognise that markets in competition law investigations 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.

2.7 **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. This leads to the research question: can consumer characteristics assist in market definition and what tools can be used to compare the consumer characteristics of different products?

This chapter answers both parts in the affirmative. Firstly, the characteristics of consumers affect their substitution choices. If market definition is concerned with studying substitutability, a study of consumer characteristics data can therefore be useful. Secondly, cluster analysis represents a useful tool for this purpose. It is useful both in the conventional sense, i.e. for delineating markets for market share calculations, and 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.

Chapter 3

Advances in Price-Time-Series Tests for Market Definition

Price elasticity estimates are the preferred tools for market definition in competition law investigations, as these estimates provide direct evidence of market power. However, Chapter 1 shows that data and time constraints in competition investigations frequently prevent the estimation of price elasticity and force practitioners to rely on less sophisticated tools. Consequently, Chapter 2 develops a complementary tool in the form of cluster analysis based on consumer characteristics. This chapter explores another set of alternative tools in the form of time-series tests of price co-movement, based on the law of one price (LOOP).

The literature has developed a range of price co-movement tests for market definition, ranging from correlation statistics to cointegration analysis. However, while extensively applied, price time series tests continue to be criticised. This chapter investigates the merits of these criticisms by considering a number of recent econometric procedures that may help to address some of the concerns raised in the literature. Specifically, the chapter considers the following research question: are the various price-time-series tests for market definition consistent and can they be improved to reflect recent time-series developments?

To this end, the chapter discusses each price test in detail, considering its focus and particular limitations, and introduces a set of new time-series tests, based on the econometric developments of the past decade. Thereafter, an empirical demonstration is attempted, using data from a recently concluded investigation into alleged anti-competitive practices in the South African milk industry at the producer-processor level. Before the extensive technical discussion of the various price tests, this chapter first considers and responds to various conceptual and empirical criticisms of price tests for market definition.

3.1 Market definition and the law of one price

Stigler and Sherwin (1985) were among the first to formally investigate the use of the LOOP in delineating markets for competition policy purposes. The LOOP, articulated as early as the late nineteenth century (see Cournot (1927) and Marshall (1920)), posits that prices within a single market should converge, allowing for some variability related to transport costs and exchange rates. Authors applying the LOOP generally argue for relative price convergence rather than for the more stringent absolute price convergence as the condition for market singularity: prices in two regions of the same market need not be equal, but should be related (Haldrup, 2003). Furthermore, the LOOP literature emphasises that a study of relative price convergence should be sensitive to the particular context, as factors such as transaction costs, degree of product differentiation and the time required for prices to adjust to arbitrage can affect conclusions regarding relative price convergence (see Hunter (2008:68-73) for a summary of the LOOP literature). However, regardless of the type of convergence, many academics and practitioners are sceptical of the use of price tests for market definition. These critics question whether a market for competition policy purposes can be identified from price relationships alone and raise a number of empirical and conceptual criticisms.

Firstly, some competition analysts question whether price co-movement can accurately approximate price elasticity estimates. Recent work by Coe and Krause (2008) uses a mainstream product differentiation model to generate price data for three products, of which two are defined close substitutes. The authors show that a number of conventional price tests have difficulty in identifying a single market for the two close substitutes and a separate market for the third. Secondly, beyond the empirical critique, other analysts criticise price tests on conceptual grounds: these critics argue that price co-movement tests only establish whether price series in different regions are ‘linked’, but do not verify whether firms have the capacity to raise prices (Massey, 2000: 317-318). Consequently, price elasticity is argued to be the only appropriate measure for market definition in competition law investigations. Hosken and Taylor (2004), Genesove (2004), and Massey (2000) argue that geographic market definition based on price co-movement, under very general conditions, could be misleading and that it “requires the economist to have substantial institutional knowledge of the markets studied” (Hosken and Taylor, 2004: 466). Along similar lines, McCarthy and Thomas (2003: 15) point to cases where two regions exhibit significant price co-movement, but supply constraints prevent producers in one region from competing with producers in the other region. They argue alternatively that regions for which price co-movement is not substantial may very well constitute a single market, if one of the regions holds excess production capacity.

This chapter investigates the conceptual and empirical criticisms levelled at price-time-series tests for market definition. Consider the conceptual criticism that price convergence tests are not suited to identifying markets in competition law investigations. The argument that price tests cannot be used to assess whether firms have the power to change prices sets too high a standard for a market definition tool. A market in competition policy is, in line with the HM test, the smallest possible space containing the smallest possible set of close substitutes that would, if owned by a single firm, enable that firm to exercise pricing power. A market definition tool is therefore aimed at ranking substitutes and identifying from this ranking the closest substitutes. The price series of two products that are close substitutes are likely to be related. In fact, one may even argue that the LOOP must hold for a set of substitutes if these substitutes are to constitute an antitrust market (Hunter, 2008). At worst, conventional price tests (read together with other relevant evidence) may identify a larger market than necessary. However, as argued in this chapter, price tests can be extended from simply being tests for the existence of relationships to also being tests for the sizes of relationships.

Consider the criticism of price tests being unable to discriminate empirically between close and less close substitutes. In their paper on the fallibility of price tests for market definition purposes, Coe and Krause (2008) consider a number of alternative data-generating processes (DGPs): they simulate data from models where cross-price elasticity is either lower or higher, which either include or exclude common cost shocks, or where dynamic price adjustment is either slower or faster. The authors then apply a battery of different price tests to the data simulated from each of these DGPs, claiming that price correlations may be useful under many of these DGPs given the absence of common shocks, but that more advanced econometric price tests generally produce incorrect conclusions concerning market boundaries. The specific details of different price tests are considered in the following section, but the focus here is on three responses to the technical criticisms raised by Coe and Krause (2008).

Firstly, price tests for market definition are frequently seen as confirmatory tools: in this view, price tests generate conclusive quantitative proof on market boundaries to support the intuition provided by anecdotal evidence. This is a problematic view, as it ignores the motivation for the use of price tests in the first place – the inability to formally estimate an econometric model and to obtain an estimate of cross-price elasticities. The lack of information forces the analyst to gather information and build a consistent story in a piecemeal fashion by combining different pieces of evidence. Under such an approach, price tests, rather than being confirmatory, become exploratory tools. The signal from these exploratory tools is likely to be imperfect as the tools only consider a subset of information. To require these imperfect tools to be accurate would be demanding too much. In fact, one may argue that even an econometric estimate of cross-price elasticity will produce an imperfect signal, as it is derived from a simple econometric specification imposing a number of assumptions and ignoring a large amount of relevant market information. All statistical tests involve the risk of incorrect inferences, but a framework that incorporates several pieces of evidence may minimise the risk by requiring the results of particular statistical tests to be supported by other qualitative evidence or even, where feasible, alternative statistical tests.

Secondly, while the first three sets of results in Coe and Krause seem to confirm their hypothesis, these results are based on idiosyncratic DGPs. When the DGPs are adjusted to be more realistic, the results on the ability of price tests to discriminate between close and less close substitutes are much less clear. Coe and Krause (2008) do not report the results for lower significance levels (say at 10% to 20%). It is true that lower significance levels indicate higher uncertainty about the inference, but they would at the very least indicate whether there are crude signals concerning price convergence. Intuitively, it would strike any econometrician as exceedingly odd that no statistical test is able to come close to simple contemporaneous correlation calculations in describing or identifying an economic relationship between two time series. Either the interpretation of the results is too limited (as argued above) or the specific tests used are not the appropriate tests (as argued below).

Thirdly, the econometric literature has developed a range of alternative time-series techniques that have better statistical power and size properties compared with the traditional techniques evaluated by Coe and Krause (2008). These alternative tests may help to improve price tests as tools for market definition. As argued in the empirical application, it is likely that at least some of the conclusions of Coe and Krause (2008) would be overturned if they were to be relied on, for example, newer tests for the presence of unit roots or equilibrium relationships. Furthermore, most of the improvements in price tests have been driven by the need to address the statistical problems experienced by earlier versions, but what is frequently ignored is that the newer tests sometimes ask different or more specific questions than the original tests. A battery of different tests may therefore provide different results that are nevertheless mutually consistent. It is therefore problematic to argue that, for example, unit root price tests should provide conclusions 'similar' to those of price correlations. The fact that the results are different does not indicate that they are misleading.

This chapter studies the market definition inferences produced by a variety of conventional and newer econometric tests of price time series. Specifically, the chapter addresses the gap in the empirical competition policy literature in ignoring the problems of the low power and size bias suffered by conventional unit root and cointegration tests – a particular problem facing practitioners who work with small data sets with an insufficient time span or data frequency. This chapter studies the performance of newer tests for a small data set for market definition purposes.

3.2 Tests of price co-movement

This section considers a variety of tests of price co-movement, which can be divided into two groups: one group of tests focuses on short-run relationships and another group on long-run relationships. Correlation statistics and Granger-causality tests fall into the first group, while unit roots, cointegration and the more recent autoregressive distributed lag models fall into the second group. A 'long run' relationship between two variables refers here to an equilibrium relationship between the two variables in levels. As noted below, the time horizon of the 'long run' is therefore an empirical matter to be decided on a case-by-case basis. The speed with which two variables adjust following a disturbance of the equilibrium will provide one measure of the case-specific 'long-run' time horizon.

3.2.1 Correlation statistics

Price tests for market definition can be performed on either time-series data or cross-sectional data. A frequently used tool for assessing price relationships between two products or regions is the correlation statistic. Correlation is strongly related to the extent of covariance between two variables, i.e. the extent to which one variable follows the movements in another variable. Formally, the sample correlation statistic for cross-section data is defined as follows:

$$\rho_{xy} = \frac{s_{xy}}{s_x s_y}$$

where

$$s_{xy} = \frac{1}{T-1}\sum_{t=1}^{T}\left(p_{x,t} - \bar{p}_x\right)\left(p_{y,t} - \bar{p}_y\right)$$

$$\bar{p}_x = \frac{1}{T}\sum_{t=1}^{T} p_x$$

$$\bar{p}_y = \frac{1}{T}\sum_{t=1}^{T} p_y$$

$$s_x^2 = \frac{1}{T}\sum_{t=1}^{T}\left(p_{x,t} - \bar{p}_x\right)^2$$

$$s_y^2 = \frac{1}{T}\sum_{t=1}^{T}\left(p_{y,t} - \bar{p}_y\right)^2$$

Sample correlation for cross-section data can be defined in a similar fashion. In the time-series domain, it is possible to adjust contemporaneous correlation statistics to incorporate temporal relationships in the form of lagged relationships, yielding the so-called cross-correlogram. The sample cross-correlation statistic between $p_{x,t-k}$ and $p_{y,t}$ is calculated by simply adjusting the covariance to reflect the lag:

$$s_{xy}(k) = \frac{1}{T-1}\sum_{t=1}^{T}\left(p_{x,t} - \bar{p}_x^*\right)\left(p_{y,t+k} - \bar{p}_y^*\right)$$

$$\bar{p}_x^* = \frac{1}{T-k}\sum_{t=1}^{T-k} p_{x,t}$$

$$\bar{p}_y^* = \frac{1}{T-k}\sum_{t=k}^{T} p_{y,t}$$

Sample cross-correlograms, however, have remained rare in competition policy applications, and most correlation analyses centre on contemporaneous correlations.

Correlation analysis, as it is applied in market definition, faces two major challenges of a statistical nature: firstly, it lacks an objective benchmark against which to assess its economic significance and, secondly, it is frequently applied without regard to statistical significance. Correlation analysis is a test of size, and it is difficult to decide whether a particular correlation statistic is economically meaningful (Forni, 2004: 450). For example, it is not clear whether a correlation statistic of 0.5 between two product price series suggests a meaningful substitution relationship for market definition purposes. One solution may be to compare correlation statistics between two variables with correlation statistics for products known to be substitutes (Davis and Garcés, 2010) or to similar correlation statistics in other markets where market integration is known (Bishop and Walker, 2002). Nevertheless, these benchmarks remain problematic. Furthermore, correlation analysis for market definition purposes is frequently employed without regard to statistical significance. A sample correlation statistic does not tell us whether the calculated sample value can be generalised to a larger population. Sample correlation statistics are point estimates that do not account for uncertainty, which may be especially large in small samples – a frequent occurrence in competition analysis. The ‘point versus interval estimates’ problem is of course not limited to market definition tools, and the need for sophisticated statistical inference has been one of the driving forces behind the move towards econometric analysis in competition cases (Davis and Garcés, 2010: 74-79).

A further criticism of correlation (although more broadly applicable to econometric methods) is the possibility of spurious correlation: correlation statistics are biased upward if underlying time series contain stochastic trends. It is possible to calculate the correlation statistic for data in first differences (to remove the stochastic trends), although first-differencing entails a significant loss of information. This issue is dealt with in subsequent sections.

It is difficult to arrive at a final summary of relationships, especially if the dynamic response of one price variable to another is protracted. It is for this reason that so-called Granger-causality tests have emerged as the second group of price tests commonly employed.

3.2.2 Granger-causality tests

Consider two price series $p_{1,t}$ and $p_{2,t}$. $p_{2,t}$ is said to Granger-cause $p_{1,t}$ if the past and present values of $p_{1,t}$ provide information to forecast $p_{2,t+1}$ at time t (Granger, 1969). Of course, as was noted in the subsequent literature, such information has no relation to true causality, but rather to time precedence. The use of Granger-causality price tests for market definition gained momentum in the 1980s (as summarised in Bishop and Walker (2002)). This is based on the notion that if the law of one price holds across two regions, price disturbances in one region should inevitably translate into price disturbances in another region; prices in one region or of one product should Granger-cause prices in another region or another product if the two products or regions share a common market.

Formally, Granger-causality tests involve comparing the following two equations:

$$p_{1,t} = \sum_{i=1}^{L} \alpha_{i1} p_{1,t-i} + g(z_t) + \varepsilon_{1,t}$$

$$p_{1,t} = \sum_{i=1}^{L} \alpha_{i1} p_{1,t-i} + \sum_{i=1}^{L} \beta_{i1} p_{2,t-i} + g(z_t) + \eta_{1,t}$$

Granger-causality then involves testing the restriction $\beta_{ij} = 0 \; \forall \; i$.

Granger-causality tests face three significant challenges as price tests for market definition. Firstly, they impose a one-way logic on market definition. By nature, Granger-causality tests are concerned with assessing direction of association. While practitioners have claimed that the focus is not on direction per se, it is not clear whether a two-way finding should be interpreted as stronger evidence of market integration compared with a one-way finding.

Secondly, Granger-causality tests are usually employed in a way that emphasises statistical significance at the expense of economic significance. For a market definition test, the issue of concern is not only whether there is a statistically significant relationship between two price series, but also whether two price series are *meaningfully* related to the extent that they will pass the SSNIP test (Bishop and Walker, 1998: 451-452). While the Granger-causality test is an attempt to overcome the use of arbitrary critical correlation levels when evaluating price correlation statistics, its application generally involves discarding size in favour of statistical fit.

Thirdly, Granger-causality tests focus on short-run relationships, ignoring the role of long-run relationships (Pagan, 1989). The tests require data to be stationary and, where data is found to be persistent (i.e. non-stationary), the analyst will usually run the tests on first differences of the non-stationary data. This, however, involves a significant loss of information – even if the focus is on short-run modelling, long-run parameters must be appropriately accounted for if the short-run parameters are not to be biased.

Persistence in price data has received increasing attention over the past decade, given its implications for monetary policy modelling (see Cogley and Sargent (2001) and the responses of Stock (2001) and Sims (2001)). More directly relevant to competition analysis, the price persistence literature has now shifted its attention from the macroeconomic price level towards prices at the disaggregated level. The evidence on persistence in disaggregated data is still conflicting and seems to suggest lower levels of persistence compared with aggregated persistence (Clark, 2006). Nevertheless, the practical implication for market definition tests is that persistence should be examined and accounted for in econometric models. The problem of high persistence in many price series has therefore motivated the shift in econometrics towards long-run tools based on unit root testing and cointegration analysis.

3.2.3 Unit root tests

Persistence is a problematic feature for regression modelling, as it generates spurious correlation or spurious regression. For example, correlation statistics are artificially high when two series each contain a stochastic trend**.** Where original price series are both unit root processes, the conventional response is to transform series into stationary versions, but this entails a significant loss of information. Furthermore, as noted earlier, the persistence implies that prices behave dynamically, in the sense that price adjustment in one area or one product in response to price adjustment elsewhere occurs slowly over time. Given these criticisms of price correlation tests, studies of long-run equilibrium relationships between variables have received significant attention in the market definition literature. Arguably, a focus on long-run equilibrium relationships returns price tests to the original motivation for using them for market definition, namely to investigate price *convergence*.

3.2.3.1 *Rationale*

One set of market definition tests that consider long-run equilibrium relationships are cointegration tests (Engle and Granger, 1987). Two non-stationary series are said to be co-integrated if a linear combination (known as the cointegrating relationship) of the series is stationary. Cointegration tests establish the existence of a cointegrating relationship, and affirmative proof is taken as sufficient indication of a single market for competition policy purposes. However, Forni (2004) argues that it is prudent to test whether the cointegrating relationship actually takes the form (1; -1), as such a one-to-one relationship indicates a perfectly integrated market. Finding a one-to-one relationship between prices in two regions clearly constitutes strong evidence that the two regions form a single market, but is not a necessary condition for market singularity. A more general cointegrating relationship of the form $(1;\ -\beta)$, for any real-valued β, also indicates market singularity, as prices are still related – even though prices only partially converge.

Nevertheless, a test for a co-integrating relationship of $(1;\ -1)$ has practical use in competition policy settings as such a test can be performed without a formal cointegration analysis. This follows because of the equivalence between a test for a cointegrating relationship of $(1;\ -1)$ and a test of whether the log price ratio for the two regions is stationary (Forni, 2004). Forni illustrates the equivalence by rewriting the log price ratio (r_t^{ij}) for two regions i and j as a one-to-one relationship between prices in the two regions:

$$r_t^{ij} = log\frac{P_{it}}{P_{jt}} = logP_{it} - logP_{jt} = p_{it} - p_{jt} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} [p_{it} \quad p_{jt}]$$

where

P_{it} is the seasonally adjusted price in region i at time t

P_{jt} is the seasonally adjusted price in region j at time t

Because $\begin{bmatrix} 1 \\ -1 \end{bmatrix} [p_{it} \quad p_{jt}]$ is stationary when there is a $(1;\ -1)$ cointegrating relationship, it implies that r_t^{ij} is stationary when there is a $(1;\ -1)$ cointegrating relationship. Therefore, in practical terms, a test for the existence of a $(1;\ -1)$ cointegrating relationship can be done by calculating the log price ratio for the two regions and then testing whether this ratio is stationary, using a standard unit root test. Forni argues that such a test is easier to apply in competition policy settings and has the added benefit of being invariant with respect to the use of nominal or real price data.

Hosken and Taylor (2004: 469) are less optimistic and argue that unit root tests can be misleading tools for market definition in two settings:

(i) In a context where a single shock is common to both price series, a unit root test is unreliable. This problem is not unique to the unit root test, but has also been studied in the context of correlation analysis (McDermott and Scott, 2000). Therefore, Forni (2004) and Boshoff (2007) argue that it is important to match unit root test results with other pieces of qualitative and quantitative evidence, including evidence on common input costs and shocks.

(ii) In a context where both price series are stationary. As noted in the following section, the ARDL test for relationships between variables in levels deals with this particular case, as it allows testing for the existence of a relationship regardless of the order of integration.

Boshoff (2007) summarises these arguments in a table showing the two outcomes of the unit root test and the associated market definition conclusion. Table 6, below, reports this table:

Table 6: Market definition using unit root test on the log price ratios for two regions

Outcome	**Geographic market definition**
Do not reject unit root	Two regions constitute separate markets
Reject unit root	Two regions constitute a single market, provided that: • At least one region has price series containing a unit root • Regions have not experienced common shocks

Source: Boshoff (2007)

Hunter (2008) and Hosken and Taylor (2004) also argue that product differentiation may bias unit root tests for market definition – a criticism that is particularly relevant to product market definition. However, one may consider a similar criticism in a geographic market context. Suppose one small and one large geographic area form a single market for a particular good: despite market singularity, a large price movement in the small area may not have a material effect on the price of the same good in the large adjacent area.

As noted, Coe and Krause (2008) recently criticised, among others, unit root tests as tools for market definition. These authors argue that unit root tests provide the incorrect market definition signal when applied to simulated price data for three products generated from an economic model in which two of the products are close substitutes. However, as is argued below, the Coe and Krause results do not necessarily tell us something about the usefulness of price tests and may be an illustration of the low power of conventional unit root tests – an issue that can easily be ameliorated by using improvements suggested by the econometrics literature over the last decade, as discussed below.

Nevertheless, it is important to acknowledge that unit root tests may be limited for practical market definition purposes as the available price series in competition investigations are frequently of a fairly short time span, rarely more than five years. A substantial literature, initiated by Perron (see Perron (1991)), shows that test statistics for unit roots are consistent only when the time span increases with the number of observations (Maddala and Kim, 1998). Therefore, even if one obtains price series of a high frequency (say monthly), unit root tests may still be problematic. Unit root tests generally require enough data points over which mean reversion could occur in order to distinguish between series with and without stochastic trends. Nevertheless, there is no absolute minimum, as mean reversion and behaviour may depend on the particular market, and it is still possible to achieve a sufficient amount of mean reversion over a relatively short span.

3.2.3.2 *Conventional unit root tests and their limitations*

Detecting persistence econometrically is difficult, as conventional OLS regression estimation and inference is biased towards suggesting statistically significant relationships for non-stationary data (Yule, 1926; Granger and Newbold, 1974). Consequently, pretesting for the existence of stochastic trends has become mandatory. However, given the bias in OLS, pretesting may itself be problematic. In response to this problem, Dickey and Fuller (1979) derived alternative asymptotic distributions under the null hypothesis of a unit root. These tests, and subsequent improvements to better account for remaining serial correlation in the residuals, have become industry-standard and are frequently employed[8].

Competition policy practitioners employing stationarity tests usually rely on Dickey-Fuller (DF) tests. Some authors even introduce the test as if it were the only test available for non-stationary (see, for example, Davis and Garcés (2010: 182). Forni (2004) employs both the DF and the so-called KPSS tests (Kwiatkowski, Phillips, Schmidt and Shin, 1992). Coe and Krause (2008: 989) also rely primarily on DF tests, although they consider alternative unit root tests in an unpublished appendix. However, Hosken and Taylor (2004) criticise the use of the DF and KPSS tests for market definition purposes, highlighting that these tests suffer from size and power problems, especially in the small samples usually encountered in competition law investigations. Size distortion refers here to the deviation of actual size (Type I error) from the nominal size selected upfront (usually 5% or 10%). Power refers to the probability of rejecting a false null hypothesis (equal to 1 minus the probability of a Type II error). The Hosken and Taylor critique accords with the pessimistic assessment of conventional unit root tests in the broader time-series literature: “Although often used, the DF [and] ADF … tests lack power against meaningful alternatives and should not be used any more” (Maddala and Kim, 1998: 92).

[8] As far as the choice of null hypothesis is concerned, the unit root literature has developed further for tests of the null of a unit root against a stationary alternative. An alternative set of tests take stationarity as the null hypothesis. For market definition purposes, the use of a non-stationary null is important in that it assumes separate markets that are tested against integration – an approach consistent with the SSNIP test that seeks to start with narrow markets and then to add to these.

The main conclusion from the unit root test literature is that the presence of residual serial correlation creates significant size distortions in smaller samples (see the summary in Haldrup and Jansson (2006)). The Said and Dickey (1984) test is an early attempt at modifying the DF test statistic by removing serial correlation using an autoregression of high order. The problem with these earlier attempts is that this improvement in test size comes at the expense of test power (Ng and Perron, 1995; Lopez, 1997). The choice of lag length therefore becomes critically important to unit root testing.

Conventional lag selection criteria for unit roots satisfy the following general form (Haldrup and Jansson, 2006):

$$IC(k) = \log \tilde{\sigma}_k^2 + kC_T/T$$

Where $\tilde{\sigma}_k^2$ is a lag-dependent function of the residual variance from the standard DF regression.

In the case of the Akaike Information Criterion (AIC), C_T is set to 2 and for the Schwartz Information Criterion (SIC), it is set to $\log T$. The AIC and SIC generally select too short a lag length, generating residual serial correlation. Ng and Perron (2001) therefore investigated the use of a modified Akaike information criterion which penalises unit root series:

$$MAIC(k) = \log \breve{\sigma}_k^2 + 2(\tau_T(k) + k)/(T - k_{max})$$

The changes to the two terms are data-based, with $\breve{\sigma}_k^2$ and $\tau_T(k)$ being more complex functions of the residual variance of the DF equation. The idea behind the adjustment is to force the IC to select a k between 0 and k_{max}, where $k_{max} = o(T)$.[9]

A set of tests that have been shown to have superior size and power properties are the modified versions of the Phillips-Perron (PP), Bhargava (B) and Elliott-Rothenberg-Stock (ERS) unit root tests proposed by Ng and Perron. Each represents an attempt at addressing the size distortion and power problems of conventional unit root tests, using the idea that time-series convergence under the null hypothesis is different from convergence under the alternative hypothesis of stationarity. These tests are discussed in the following subsection.

3.2.3.3 *Improvements in unit root tests*

[9] $k_{max} = o(T)$ means that T grows much quicker than k_{max}.

Ng and Perron (1995) develop three so-called M-tests. The M-tests are the modified Phillips-Perron (PP) tests with improved power properties, relying on the property that time-series convergence under the unit root null hypothesis is different from convergence under the alternative hypothesis of stationarity.

The first test statistic is:

$$MZ_\rho = Z_\rho + \frac{T}{2}(\hat{\rho} - 1)^2$$

Perron and Ng (1996) show that MZ_ρ converges in distribution to the original Dickey and Fuller (1979) distribution, given that $\hat{\rho} - 1 = O_p(T^{-1})$ under the null hypothesis of a unit root[10].

The second test statistic is a modified version of Bhargava (1986), which in turn, was an improvement on Sargan and Bhargava (1983):

$$MSB = \sqrt{\frac{1}{\varpi_{AR}^2}\frac{1}{T^2}\sum_{t=2}^{T} p_{t-1}^2}$$

Given that $\sum_{t=1}^{T} p_t^2$ is $O_p(T^2)$ if p_t is $I(1)$ and $O_p(T)$ if p_t is $I(0)$, the statistic is $O_p(T^{-1})$ under the alternative hypothesis of stationarity, implying that the test statistic is bounded by zero from below.

There is a close relationship between MSB and the conventional $PP\ Z_t$ test statistic:

$$Z_t = MSB \cdot Z_\rho$$

Perron and Ng (1996) use this to derive a third test statistic MZ_t:

$$MZ_t = Z_t + \frac{1}{2}MSB(\hat{\rho} - 1)^2$$

As is the case for MZ_ρ, the second term of the MZ_t statistic is asymptotically zero. The important aspect of the M-tests is their improved size properties *provided* the particular autoregressive spectral density estimator suggested by Perron and Ng (1996) is used. Alternative kernel-based estimators do not produce the same size improvements.

It is not only small samples that generate size problems for the application of conventional unit root tests in competition law investigations. Frequent measurement errors and outliers in corporate datasets also

---

[10] $O_p(T^{-1})$ means that the series tends to zero.

significantly bias conventional test sizes, and Vogelsang (1999) demonstrates that the size of M-tests is robust in terms of these problems (Haldrup, Montanés and Sanso, 2005; Haldrup and Jansson, 2006). From a practitioner perspective, the tests are also easy to apply, as they are available in a number of mainstream econometrics software packages, including Eviews.

It turns out that the estimator of the long-run variance, $\hat{\varpi}^2$, has significant implications for the power and size properties of the M-tests. Most unit root tests rely on kernel-based estimators, but Perron and Ng (1996) question their use and show that these estimators may aggravate the size distortion[11]. As an alternative, Perron and Ng suggest an autoregressive spectral density estimator:

$$\hat{\varpi}_{AR}^2 = \frac{\tilde{\sigma}_k^2}{(1 - \sum_{j=1}^{k-1} \tilde{\gamma}_j)^2}$$

where k is based on an appropriate information criterion and $\tilde{\gamma}_j$ is the coefficient of the jth lagged, differences series in the DF equation. Although the use of the spectral density estimator significantly enhances the size properties of even conventional unit root tests, it does not remove the size distortion. Perron and Ng show that it is only when employed with the M-tests that it ensures that size distortion is removed.

Deterministic components play an important role in the power of unit root tests. Excluding a deterministic linear trend, for example, leads to zero asymptotic power for conventional DF tests (Haldrup and Jansson, 2006). Elliott, Rothenberg and Stock (1996) suggest an improved detrending technique, based on local generalised least squares (GLS) methods. Ng and Perron (2001) improve on this technique and demonstrate that the M-tests based on GLS detrended data are uniformly more powerful than M-tests based on data detrended with other techniques. They also argue that asymptotic power is close to the power of the so-called point-optimal unit root tests suggested by Elliott, Rothenberg and Stock (1996). To test the relation between these unit root tests, Ng and Perron (2001) also develop a fourth modified test based on the feasible point-optimal tests of Elliott, Rothenberg and Stock (1996), such that the modified test has the same asymptotic distribution as the original. The MP_T^{GLS} test statistic, based on GLS-detrended data $\tilde{p}$ is as follows:

$$MP_T^{GLS} = \frac{1}{\varpi_{AR}^2}(c^2 \frac{1}{T^2} \sum_{t=2}^{T} \tilde{p}_{t-1}^2 - (1-c)\frac{1}{T}\tilde{p}_T^2)$$

[11] This follows from the relationship between kernel-based estimator $\hat{\varpi}^2$ and $\hat{\rho}$ – the latter experiences size distortion in small samples under the null of unit root, resulting in inefficiency for $\hat{\varpi}^2$.

where c is a function of the GLS detrending procedure and input parameters.

This chapter does not explicitly study the global power of these unit root tests. The literature focuses on *local* asymptotic power, i.e. the performance of the test statistic in "a neighbourhood of the null hypothesis" (McManus, 1991). In contrast, *global* asymptotic power refers to a case where the data generating process (DGP) under the alternative hypothesis is constant, and the power envelope is calculated over increasing sample sizes. However, it may be that the DGP under the alternative hypothesis approaches the null DGP over larger sample sizes, resulting in asymptotic power not converging to unity (Neyman, 1937). Seo (2005) highlights this issue and Perron and Qu (2007) suggest an adjustment to the original Ng and Perron (2001) procedure that would address this problem. Local power receives most of the attention in the unit root literature, including the recent comprehensive survey of power envelopes by Haldrup and Jansson (2006).

3.2.4 ARDL model and bounds test for level relationships

The choice concerning which price test is the more appropriate for market definition depends on whether market definition considers only *whether* prices are related or also *to what extent* prices are related. Unit root tests are strict tests for both the existence of a relationship and, if present, for a one-to-one long-run relationship. As noted earlier, unit root tests do not consider long-run relationships that are not one-to-one. Cointegration analysis addresses this shortcoming, allowing the analyst to test for a general set of long-run relationships.

Cointegration analysis on price series and unit root tests on price ratios both assume that what matters for market definition is *long-run* equilibrium relationships. However, as Bishop and Walker (2002) note, market definition, via the SSNIP methodology, is concerned with the response to a short- to medium-run price change. In the long run, consumers will have the capacity to respond to price changes: relying on long-run equilibrium is therefore likely to overstate the size of the market. Of course, it may still be possible to argue that cointegration can assist in building an error-correction model (ECM), by adding long-run relations (represented in error-correction term z_{t-1}) and enabling a more accurate estimation of the short-run dynamics (represented as $lagged(\Delta p_{1,t}, \Delta p_{2,t})$):

$$\Delta p_{1,t} = \gamma_1 z_{t-1} + lagged(\Delta p_{1,t}, \Delta p_{2,t}) + \varepsilon_{1,t}$$

$$\Delta p_{2,t} = \gamma_2 z_{t-1} + lagged(\Delta p_{1,t}, \Delta p_{2,t}) + \varepsilon_{2,t}$$

The coefficients in $lagged(\Delta p_{1,t}, \Delta p_{2,t})$ indicate the degree to which short-run movements in the two price series are related and can be used to define the relevant market. However, this is rarely the way in which cointegration analysis is employed in market definition. Usually, the focus falls on the error-correction term z_{t-1} taking the general form. It is of course still possible to focus on the coefficients γ_1 and γ_2 to infer the speed of adjustment – and, hence, the 'horizon' of the relationship between the price series. If the adjustment is quick (i.e. if the coefficients are large), the relationship is likely to be a short-run relationship. However, the focus on the error-correction term is usually limited to a test for the existence of cointegration (Werden and Froeb, 1993).

Furthermore, even if the long-run equilibrium is considered the more appropriate part for market definition purposes, the cointegration analysis is still exposed to potential unit root pretesting bias – given that one has first to verify that the price series are non-stationary prior to performing the cointegration analysis. The implications of an incorrect inference on unit root properties will be substantial, as the distribution of the test statistic is vastly different for a unit root process than it is for a near unit root process. This probably limits the use of cointegration analysis in a market definition context, given the small samples.

An alternative approach to cointegration is the bounds testing approach developed by Pesaran, Shin and Smith (2001) (PSS). Instead of testing for the presence of a cointegration relationship (and also the extent of cointegration), the aim for the bounds test is to test for the presence of a levels relationship *regardless of the order of integration of input variables*. Input variables may be $I(0)$, $I(1)$ or cointegrated. This implies that the bounds test also allows for a broader class of levels relationships to qualify for market definition purposes than allowed for by unit root tests on price ratios (see previous subsection).

The bounds test is based on a conditional error-correction model (ECM), and the test statistic is the F-test for the statistical significance of lagged variables in *levels* in the ECM. PSS derive asymptotic distributions for the test statistic that assume either all variables $I(0)$ or all variables $I(1)$. These two sets form the upper and lower bounds of the test. If the test statistic exceeds the upper bound or falls below the lower bound, the inference is straightforward, but the outcome is unknown when the test statistic falls between the two bounds. In the latter case, it is still necessary to test for stationarity.

The ARDL approach is easy to implement, which further recommends it for use in market definition. As discussed below, the procedure involves a single-equation estimation using OLS and a straightforward testing of coefficient restrictions using the F-test. Consequently, most practitioners with only an introductory technical knowledge of econometrics can perform bounds tests.

3.2.4.1 *Bounds test methodology*

Consider a bivariate $VAR(q)$ model for the price vector $= \begin{bmatrix} p_{1,t} \\ p_{2,t} \end{bmatrix}$. Assume all series are strictly $I(0)$, strictly $I(1)$ or cointegrated. The error term has zero conditional mean and is homoscedastic. The following *conditional* ECMs can be constructed from this VAR, the first conditioning on $p_{1,t}$ and the second on $p_{2,t}$:

$$\Delta p_{1,t} = \alpha_0 + \sum_{i=1}^{q} \alpha_{1,i} \Delta p_{1,t-i} + \sum_{i=0}^{q} \alpha_{2,i} \Delta p_{2,t-i} + \beta_1 p_{1,t-1} + \beta_2 p_{2,t-1} + \varepsilon_{p1,t}$$

$$\Delta p_{2,t} = \varphi_0 + \sum_{i=1}^{q} \varphi_{1,i} \Delta p_{2,t-i} + \sum_{i=0}^{q} \varphi_{2,i} \Delta p_{1,t-i} + \beta_3 p_{1,t-1} + \beta_4 p_{2,t-1} + \varepsilon_{p2,t}$$

From these models, one can test for the existence of a long-run relationship between $p_{1,t}$ and $p_{2,t}$, and if verified, estimate the long-run relationship. Formally, the structure of the null hypotheses and alternative hypotheses can be stated as follows:

$$H_0 = H_0^{\beta 1} \cap H_0^{\beta 2}$$

$$H_1 = H_1^{\beta 1} \cup H_1^{\beta 2}$$

where

$H_0^{\beta 1}: \beta_1 = 0$

$H_0^{\beta 2}: \beta_2 = 0$

$H_1^{\beta 1}: \beta_1 \neq 0$

$H_1^{\beta 2}: \beta_2 \neq 0$

The alternative hypothesis therefore also covers the degenerate cases $\beta_1 \neq 0$, $\beta_2 = 0$ and $\beta_1 = 0$, $\beta_2 \neq 0$. PSS show that, only under the assumption that $\beta_1 \neq 0$ can one derive a conditional level relationship for $p_{1,t}$ and $p_{2,t}$. If $\beta_2 = 0$, the conditional ECM clearly has no level effects, and there is no possibility of any level relationship. If $\beta_2 \neq 0$, then $\Delta p_{1,t}$ depends on the levels of $p_{2,t}$ *only* through its relation with the coefficients of $p_{2,t}$ in the original (not conditional) ECM, which does not point to a *long-run* relationship. As PSS acknowledge, there is still the potential for short-run relationships.

Null hypotheses of the above form involve multiple parameter restrictions and are usually tested using an F-statistic, which is compared with some critical value at a prescribed significance level. PSS compute asymptotic critical values for tests of this hypothesis on the conditional ECM of interest. The F-distribution for these joint hypotheses depends critically on the order of integration of the conditioning price variable. To avoid pretesting, PSS introduce a bounds testing approach relying on two critical values: simulation evidence suggests that the critical values for $I(0)$ conditioning variables form a lower bound, while critical values for the $I(1)$ case form an upper bound. To ensure that the conditional ECM is congruent with the underlying data, PSS develop separate critical values for different specifications involving deterministic components (trends and intercepts) of the VAR.

One of the contributions of this chapter is to introduce new time-series tests with improved statistical power against the null hypothesis compared with conventional time-series techniques. The bounds test offers an important innovation in overcoming the pretesting problem faced by formal cointegration analysis. The test has improved power (of correctly rejecting), but this power depends critically on the derived asymptotic distribution. As a further enhancement, it is necessary to investigate the small sample properties of the F-test used in the bounds testing approach. The critical values reported by PSS are based on sample sizes of 500 and 1000. While the assumption is that ‘long enough’ time-series are likely to ensure convergence of the F-test on these asymptotic distributions, this assumption is highly problematic for the smaller samples typically encountered in competition work. Using Monte Carlo methods and different sample sizes, Turner (2006) estimates a response surface for a particular number of variables and specific deterministic components (Cases 1 to 5), which he then compares to the asymptotic critical values. Turner (2006) shows that it is necessary to calculate sample-specific critical values for samples as large as 300. Narayan (2005) develops a set of critical value bounds specifically for Cases 2 to 5 for sample sizes ranging from 30 to 80. Narayan uses the same approach and computer code used by PSS in generating the asymptotic results. The conditional ECM results are evaluated using these finite-sample critical values.

3.2.5 Summary

The previous subsections consider a variety of price-time-series tests for market definition, and their arguments can be summarised in two points. Firstly, price tests can be divided into tests for short-run relationships (correlation statistics and Granger-causality) and tests for long-run relationships (unit root tests, cointegration analysis and ARDL bounds tests). In fact, each price test aims to evaluate a relationship by asking a question unique to that test, i.e. each test proceeds from a specific null hypothesis. Therefore, it is incorrect to argue that the divergent results of different price tests disqualify these tests as useful tools for market definition. As shown in the empirical application, different price tests provide different perspectives on price relationships. Secondly, the empirical criticisms levelled at conventional price tests, including the poor small-sample performance of the Dickey-Fuller unit root test and Johansen cointegration test, can be addressed by employing new tests that account for power and size problems in small samples. To this end, this chapter introduces various recent econometric advances, which are not widely adopted in practice.

Given these arguments, it is important to illustrate the improvement empirically by contrasting the results from conventional price tests with those from the improved tests suggested above. The following section uses data from a recent South African competition investigation for this comparison.

3.3 Case description and qualitative evidence on the relevant market

The relative performance of conventional and new price tests is illustrated in the context of geographic market definition. Specifically, the comparison is based on a recent South African competition investigation, which involved the alleged anti-competitive business practices of dairy processors in the South African milk industry.

3.3.1 Case description

The competition investigation focused on exclusive supply agreements concluded between dairy processors and dairy farmers belonging to SAMILCO, an industry body representing dairy farmer interests. The agreements required members of SAMILCO to sell all of their milk production to the processor, or risk losing membership of SAMILCO. Agreements were concluded for a three-year period, after which the agreements would continue but with the option of terminating after a six month notice period. The competition authority held that these agreements constituted anti-competitive vertical restraints. The competition investigation also covered a range of other practices, including coordination among dairy processors and price fixing, but the focus here is on the market definition question as it relates to the vertical restraints issue (see Competition Tribunal (2008: 3) for a summary of the allegations).

Given that the alleged agreements were concluded upstream between dairy processors and farmers, but may have had an effect on final downstream consumers, two product markets are relevant: there was an upstream market for fresh milk between dairy processors and dairy farmers as well as a downstream market between dairy processors and final consumers (via retailers). In this case, the important market definition question is the geographic scope of the upstream market. One of the dairy processors investigated owned processing plants across three southern milk regions (labelled as Western Cape, Southern Cape and Eastern Cape). For the purpose of assessing whether this particular processor was dominant, it is therefore necessary to establish whether the adjacent southern regions constitute a single market[12].

As argued earlier in the response to the empirical criticisms against price tests, market definition cannot be based on a single tool, but should involve an analysis of several pieces of evidence. Therefore, the following subsection summarises the non-price descriptive evidence in the milk case, including product flows and other qualitative evidence. The descriptive evidence suggests a particular geographic market, and subsequent price tests are then used to investigate econometrically whether the price relationships in this suggested market are consistent with the hypothesis of a single market (see Copenhagen Economics (2003) and McCarthy and Thomas (2003: 8-9) for a more detailed explanation of a structure that integrates descriptive and econometric evidence).

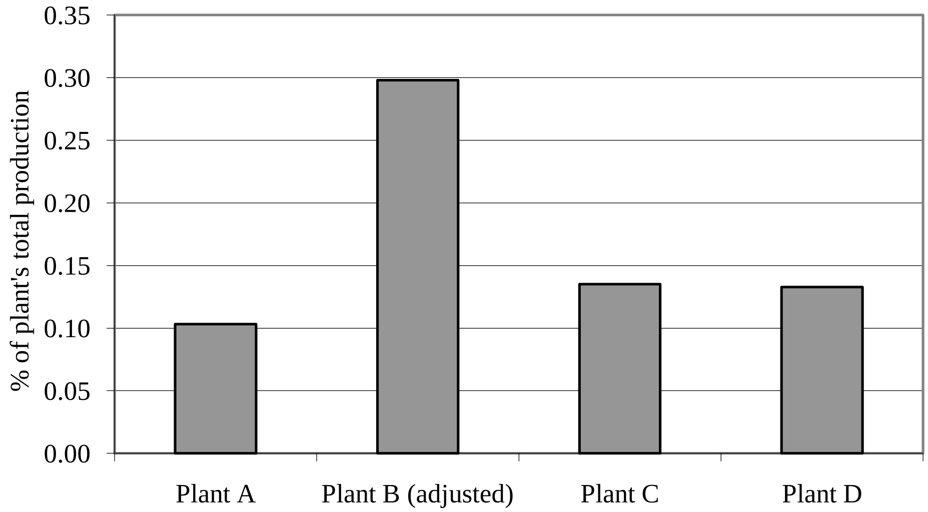
3.3.2 Descriptive evidence on geographic substitutability

One popular type of descriptive evidence frequently used in geographic market definition is evidence on product flows, as this mirrors the well-known Elzinga and Hogarty (1973) test. In this case, data on aggregate milk flows among the regions was not available. Instead, data on processed milk flows among the four different plants of the dairy processor under investigation were analysed. If these flows were not significant, it would be a first indication that transport costs were not prohibitive – and that a single market is plausible. Boshoff (2007) calculated average monthly milk inflows for each plant, expressed as a percentage of overall production. Average inflows were virtually zero (all were less than 0.5% of production) over the period 2004 to 2005, as shown in Figure 9.

However, the average calculations for Plant B exclude flow data for June and July 2004, masking significant milk transfers (of 24% and 13% of overall production) to Plant B during those months. While product flows among the four southern plants did not occur frequently, it appears that transport costs occasionally permitted large transfers.

[12] In the current case, the focus falls on the use of buyer power. The HM test is articulated in terms of *seller* power (in other words, the market is defined by considering the response of the processor to a SSNIP by the dairy farmer). However, it is also possible to see the price tests as an implementation of an SSNDP (small but significant non-transitory decrease in price) by a buyer.

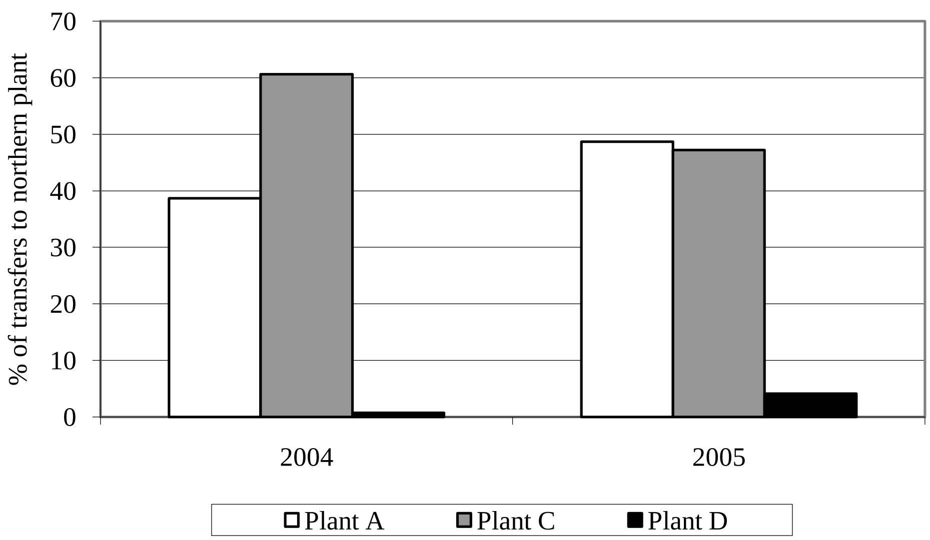
Figure 9: Average monthly imports from other regions as percentage of production at different plants



Source: Boshoff (2007)

Stronger evidence for a single market emerges when one considers product flows between the southern milk plants and the single northern plant of the processor. Since June 2003, the dairy processor sourced milk for the northern plant from farms in the south – following disagreement between the processor and farmers close to the northern plant regarding milk prices. The dairy processor also sourced milk from other plants owned by competitors, which formed part of the competition investigation. Figure 10 offers a breakdown of the sources of the southern milk for the northern plant for 2004 to 2005:

Figure 10: Share of selected southern plants in annual milk transfers to the northern plant



Sources are labelled by their collection point (i.e. the specific southern plants from which they were despatched to the north). Plants A and C (situated furthest from the northern plant) dominated milk transfers to the north, which strongly suggests that transport costs were low enough to support a single market among the southern plants or, even, a national market that also included northern processors.

Similar evidence was also provided for other processors, including media reports on the new transport vehicles used by a dairy processor to ship milk from KwaZulu-Natal to other parts of the country (see, for example, Mawson (2005)). Descriptive and anecdotal evidence on product flows therefore strongly suggested a single geographic market encompassing a number of southern plants and, perhaps, even plants in the north. The following section further explores the proposed single market hypothesis, applying the various price tests discussed earlier to milk price data obtained from the processor under investigation.

3.4 Price test results

3.4.1 Milk price data

Price tests require representative milk price data. Prices negotiated between different processors and farmers differed significantly and, ideally, it would have been useful to have had access to all of these individual prices in order to do a fine-grained analysis of price relationships. However, in this case such data was not available. Fortunately, data from an agricultural co-operative, the Southern Africa Milk Co-operative (SAMILCO), was available. SAMILCO represents a large portion of dairy farmers in the southern regions and calculates an average monthly milk price for each of the Western Cape[13], Southern Cape and Eastern Cape regions. This data is used for the subsequent price tests.

Haldrup (2003: 16) argues that competition practitioners should account for seasonal features in prices, as competition cases frequently involve high-frequency data of a short span. Consequently, the X12 seasonal adjustment technique is used to seasonally adjust the three price time-series. Seasonal adjustment using alternative procedures, including the TRAMO/SEATS procedure, did not produce significantly different results – consistent with findings in the literature (see the summary in Fok, Franses and Paap (2006)).

Hosken and Taylor (2004) emphasise that the competition analyst should not lose sight of institutional details that may alter a conclusion based on price tests. This is important in the case of the relationship between the Western Cape and Southern Cape prices. Before 2005, the market relations between dairy processors and farmers were governed by supply agreements, which influence prices, but prices in

[13] The Western Cape region does not refer to the Western Cape Province.

different regions were not contractually linked. Nevertheless, market prices in the different regions generally co-moved until January 2004, when the prices of Western Cape milk increased significantly following very strong actual and expected increases in milk demand. Prices also rose in the Southern and Eastern Cape but to a lesser extent. Southern Cape farmers therefore demanded higher prices from dairy processors, failing which farmers signalled that they would use the arbitrage opportunity to sell their product in the Western Cape. This resulted in a determined effort to match prices in the two regions from the start of 2005, which culminated in a renewed agreement in December 2005 between the processor under investigation and its Southern Cape suppliers. This agreement explicitly linked the average milk price in the Western Cape with the average milk price in the Southern Cape, allowing for transport costs between the two regions. One could argue that the 2005 changes altered the market dynamics and artificially created a single market across the two regions. Alternatively one could argue that these changes are a reflection of interaction in a market that has always included both regions. Either way, it is necessary to study two sample periods, one including and one excluding the 2005 data. The need for separate sample periods is also important in the relationships involving the Eastern Cape, as price ratios between the Eastern Cape and other regions are also affected.

3.4.2 Time-series properties of milk prices

As argued earlier, the properties of time-series matter greatly for the application of econometric tests of price co-movement.

Figure 11: SAMILCO prices in Western Cape, Southern Cape and Eastern Cape, January 2002 – December 2005

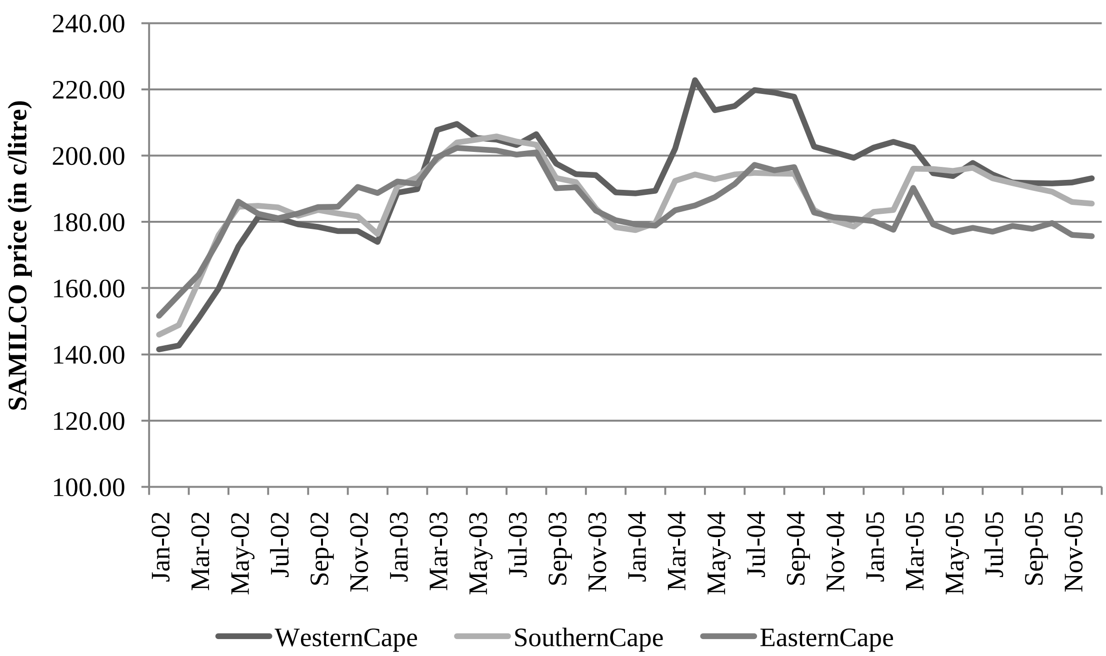


Figure 12: Correlograms for milk prices, January 2002 – December 2005

(a) Western Cape

Date: 05/24/10 Time: 08:12
Sample: 2002M01 2005M12
Included observations: 48

Lag	AC	PAC	Q-Stat	Prob
1	0.863	0.863	38.072	0.000
2	0.723	-0.087	65.376	0.000
3	0.584	-0.080	83.549	0.000
4	0.471	0.021	95.674	0.000
5	0.398	0.073	104.50	0.000
6	0.330	-0.037	110.74	0.000
7	0.268	-0.033	114.94	0.000
8	0.209	-0.020	117.54	0.000
9	0.150	-0.033	118.92	0.000
10	0.093	-0.040	119.47	0.000
11	0.019	-0.120	119.49	0.000
12	-0.048	-0.034	119.64	0.000

(b) Southern Cape

Date: 05/24/10 Time: 08:12
Sample: 2002M01 2005M12
Included observations: 48

Lag	AC	PAC	Q-Stat	Prob
1	0.802	0.802	32.841	0.000
2	0.604	-0.109	51.898	0.000
3	0.427	-0.066	61.619	0.000
4	0.324	0.088	67.351	0.000
5	0.266	0.040	71.298	0.000
6	0.192	-0.096	73.405	0.000
7	0.099	-0.090	73.981	0.000
8	0.002	-0.063	73.981	0.000
9	-0.057	0.012	74.179	0.000
10	-0.149	-0.198	75.587	0.000
11	-0.211	-0.023	78.466	0.000
12	-0.281	-0.102	83.743	0.000

(c) Eastern Cape

Date: 05/24/10 Time: 08:13
Sample: 2002M01 2005M12
Included observations: 48

Lag	AC	PAC	Q-Stat	Prob
1	0.790	0.790	31.868	0.000
2	0.642	0.047	53.364	0.000
3	0.462	-0.155	64.761	0.000
4	0.341	0.019	71.096	0.000
5	0.284	0.119	75.597	0.000
6	0.221	-0.049	78.380	0.000
7	0.087	-0.272	78.820	0.000
8	-0.055	-0.149	78.999	0.000
9	-0.155	0.045	80.475	0.000
10	-0.254	-0.123	84.551	0.000
11	-0.265	0.014	89.093	0.000
12	-0.292	-0.060	94.790	0.000

Figure 11 suggests changes in the distributional properties of the individual SAMILCO price series, with the series moving arbitrarily far from a deterministic trend. The graphical intuition is confirmed by the autocorrelogram and, in particular, the partial autocorrelogram in Figure 12 – suggesting the presence of unit roots in the different price series. Table 7 presents the Dickey-Fuller unit root tests discussed earlier applied to the individual SAMILCO price series for the Western Cape, Southern Cape and Eastern Cape, for a range of lag lengths. Table 8 repeats the analysis, using the Ng-Perron tests. Both tables confirm that the null hypothesis of a unit root cannot be rejected for any of the three series. However, the Dickey-Fuller test results for the Southern Cape do appear to be unclear, with evidence at smaller lag orders indicating stationarity. This uncertainty is not present in the Ng-Perron test results.

Table 7: Augmented Dickey-Fuller (1979) tests on milk prices, January 2002 – December 2005

Lag	***t*-stat**	**p-value**
Western Cape		
1	-2.62	0.27
2	-2.99	0.15
3	-2.84	0.19
4	-2.17	0.49
5	-1.96	0.61
6	-2.00	0.58
Southern Cape		
1	-4.04	0.01***
2	-4.50	0.00***
3	-3.32	0.08*
4	-1.99	0.59
5	-2.11	0.53
6	-1.99	0.59
Eastern Cape		
1	-3.18	0.10*
2	-3.53	0.05
3	-2.86	0.18
4	-1.81	0.68
5	-2.60	0.28
6	-3.37	0.07*

Note: *** Reject at 1%, ** Reject at 5%, * Reject at 10%

Table 8: Ng-Perron (1996, 2001) tests on milk prices, January 2002 – December 2005

Lag	MZ_a	MZ_t	MSB	MP_T
		Western Cape		
1	-3.72	-1.30	0.35	23.58
2	-4.30	-1.41	0.33	20.68
3	-4.31	-1.41	0.33	20.62
4	-4.55	-1.45	0.32	19.59
5	-7.58	-1.90	0.25	12.12
6	-6.48	-1.75	0.27	14.08
		Southern Cape		
1	-3.33	-1.27	0.38	26.97
2	-3.91	-1.38	0.35	23.05
3	-2.93	-1.19	0.41	30.51
4	-2.10	-1.00	0.48	42.00
5	-2.75	-1.15	0.42	32.38
6	-4.40	-1.47	0.33	20.54
		Eastern Cape		
1	-2.80	-1.13	0.40	30.88
2	-4.60	-1.47	0.32	19.49
3	-4.20	-1.40	0.33	21.24
4	-2.16	-0.98	0.45	38.93
5	-2.58	-1.08	0.42	33.27
6	-11.71	-2.39	0.20	7.93

Note: *** Reject at 1%, ** Reject at 5%, * Reject at 10%

3.4.3 Price correlation analysis

Table 9 and Table 10 report the contemporaneous correlation among the three price series (seasonally adjusted) separately for the sample period including and excluding 2005. As argued earlier, price tests, including simple correlation calculations, must account for the presence of unit roots. Therefore, Table 9 focuses on levels, whereas Table 10 focuses on first differences. Table 10 suggests that correlation is reduced once the unit root is removed by first-differencing. However, even in Table 10, the correlation is about 0.7 between Eastern and Southern Cape prices and around 0.5 for Western Cape prices, on the one hand, and Southern or Eastern Cape prices, on the other.

Table 9: Correlation matrix of milk prices in levels

(a) January 2002 – December 2004

	Western Cape	Southern Cape	Eastern Cape
Western Cape	1.00	0.82	0.74
Southern Cape	0.82	1.00	0.95
Eastern Cape	0.74	0.95	1.00

(b) January 2002 – December 2005

	Western Cape	Southern Cape	Eastern Cape
Western Cape	1.00	0.82	0.67
Southern Cape	0.82	1.00	0.81
Eastern Cape	0.67	0.81	1.00

Table 10: Correlation matrix of milk prices in first differences

(a) January 2002 – December 2004

	Western Cape	Southern Cape	Eastern Cape
Western Cape	1.00	0.51	0.48
Southern Cape	0.51	1.00	0.75
Eastern Cape	0.48	0.75	1.00

(b) January 2002 – December 2005

	Western Cape	Southern Cape	Eastern Cape
Western Cape	1.00	0.50	0.45
Southern Cape	0.50	1.00	0.72
Eastern Cape	0.45	0.72	1.00

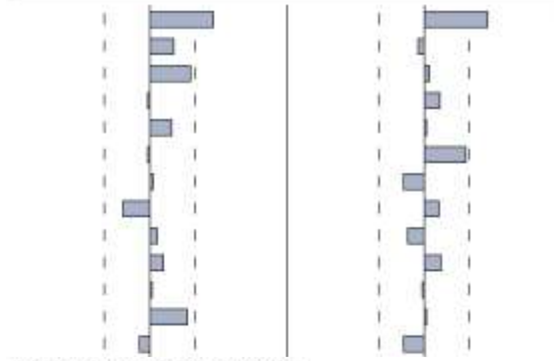
As argued earlier, correlation calculations based on first-differenced variables run into a number of limitations, including the fact that they ignore the longer-run information in the two variables. Before moving to more advanced techniques that address this problem, it is useful to consider lagged correlations – price changes in one region are unlikely to be transmitted instantaneously to another region.

Figure 13: Cross-correlograms of milk prices in first differences, January 2002 – December 2004

(a) Western and Eastern Cape:

Date: 05/24/10 Time: 07:57
Sample: 2002M01 2004M12
Included observations: 35
Correlations are asymptotically consistent approximations

WCAPE,ECAPE(-i) WCAPE,ECAPE(+i)

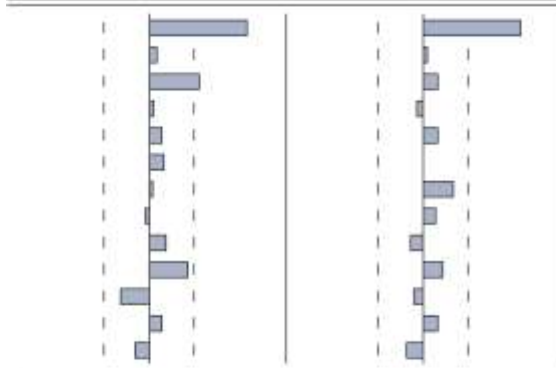


i	lag	lead
0	0.4804	0.4804
1	0.1823	-0.0454
2	0.3053	0.0427
3	-0.0140	0.1167
4	0.1680	0.0181
5	-0.0151	0.3147
6	0.0318	-0.1528
7	-0.2040	0.1132
8	0.0614	-0.1201
9	0.1048	0.1309
10	0.0243	-0.0148
11	0.2815	0.0191
12	-0.0796	-0.1543

(b) Southern and Eastern Cape:

Date: 05/24/10 Time: 08:03
Sample: 2002M01 2004M12
Included observations: 35
Correlations are asymptotically consistent approximations

SCAPE,ECAPE(-i) SCAPE,ECAPE(+i)

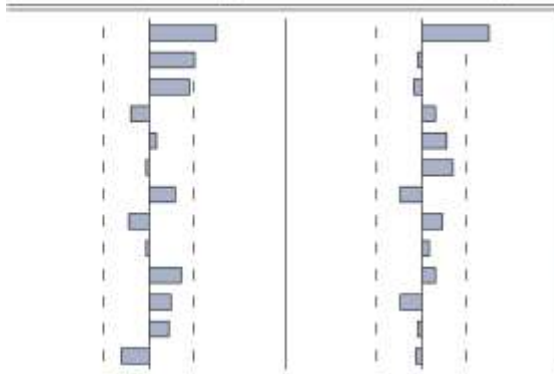


i	lag	lead
0	0.7453	0.7453
1	0.0585	0.0345
2	0.3797	0.1109
3	0.0385	-0.0434
4	0.0919	0.1085
5	0.1111	0.0071
6	0.0309	0.2338
7	-0.0262	0.0925
8	0.1305	-0.0902
9	0.2901	0.1520
10	-0.2138	-0.0718
11	0.0980	0.1067
12	-0.0985	-0.1172

(c) Western and Southern Cape:

Date: 05/24/10 Time: 09:06
Sample: 2002M01 2004M12
Included observations: 35
Correlations are asymptotically consistent approximations

WCAPE,SCAPE(-i) WCAPE,SCAPE(+i)



i	lag	lead
0	0.5116	0.5116
1	0.3446	-0.0249
2	0.3081	-0.0624
3	-0.1333	0.1046
4	0.0575	0.1890
5	-0.0235	0.2401
6	0.2001	-0.1579
7	-0.1538	0.1620
8	-0.0241	0.0609
9	0.2502	0.1052
10	0.1670	-0.1605
11	0.1535	-0.0305
12	-0.2126	-0.0455

Figure 13 reports results for the sample period January 2002 – December 2004. The cross-correlogram reports both lagged and leading correlations from lag zero (i.e. contemporaneous correlation) to lag 12 and also indicates, via two broken lines, the 95% confidence level for the correlations[14]. It is striking that the size of cross-correlations never exceeds 0.4 for any of the price pairs. Also, whereas the correlation between Southern and Eastern Cape is strongest contemporaneously but then declines sharply, the lagged correlation for the Western and Southern Cape declines more gradually from zero to two lags. All of the cross-correlograms suggest that correlations for lags longer than two quarters are not significantly different from zero at a 95% confidence level. This may suggest that the correlation evidence for market integration is less strong when lagged relationships are accounted for, although it is still not clear what the benchmark is.

If one relies on contemporaneous correlation statistics to define markets, one would conclude that the evidence supports a single market across the Southern and Eastern Cape with less strong support for also including the Western Cape. There are three potential problems with this interpretation. First, the conclusion is arbitrary in the absence of a benchmark for deciding whether the correlation is statistically or economically significant for market definition purposes. Second, if conclusions based on lagged correlations are preferred to contemporaneous correlations, one may conclude in favour of separate markets given that the lagged correlations are much smaller. Third, the market delineation is based on short-term relationships only and does not account for long-term dynamic relationships.

3.4.4 Lag length selection

The limits of static correlation analysis have motivated the development of dynamic time-series econometrics. Econometric techniques, including the Granger-causality tests, unit root tests and the ARDL models, require specification of the lag length. From a practical perspective, lag order selection involves a trade-off between parsimony and fit: larger lag orders may remove residual serial correlation but increase the complexity of a model, which is especially problematic when data constraints are high – as they usually are in competition cases.

The econometric literature has developed a number of alternative test criteria for selecting the optimal lag length. Where the tests results conflict (by proposing different lag lengths), the interpretations of the majority of the tests are adopted. The results of these tests for the milk data are reported fully in Appendix

[14] The confidence interval for the correlation between y_t and x_{t-s} is derived from the regression $y_t = \beta_0 + \beta_1 x_{t-s}$, on which the hypothesis $\beta_1 = 0$ is tested using a likelihood ratio (LR) test. The properties of the LR test are used to derive a distribution for the cross-correlation statistic.

C, and Table 11 summarises the optimal lag length suggested:

Table 11: Optimal lag lengths

System	**Optimal lag length**
Western & Eastern Cape	1
Southern & Eastern Cape	3
Western & Southern Cape	1

Lag selection tests may be misleading in the current context. The milk price dataset has a short sample period, which renders econometric results for specific lags quite sensitive to one or two exceptional data points. This is particularly evident in the case of the test for a lag of six months for the Western and Eastern Cape model: all of the results but the likelihood ratio (LR) test suggest a lag length of one month. However, the LR test suggests six months. This is possibly due to the econometric model picking up information related to the sixth lag, as is evident from the cross-correlogram result for the same lag length, presented earlier. However, this does not appear to be a systematic feature of the data, given the absence of significant lags between three and six months.

The optimal lag lengths are therefore indicative only, and the following econometric tests consider a broader range of lag lengths to ensure that the conclusions are robust in terms of errors in lag length selection.

3.4.5 Granger-causality analysis of prices

As discussed previously, a well-known technique for studying dynamic bivariate relationships is the so-called Granger-causality test. As argued, the test also requires stationary data and therefore suffers from the same loss of information as the correlation statistic, but it has the advantage of accounting for *dynamic* relationships. Tables 8 to 10 report Granger-causality tests for both sample periods (2002-2004 and 2002-2005) for the three bivariate relationships. The tables report the size of the F-test statistic as well as the associated p-values. Results for the optimal lag length are highlighted in boldface.

Table 12 finds some evidence of a Granger-causal relationship between Eastern Cape and Western Cape milk prices in the sample period excluding 2005, but no results for the sample period including 2005. Table 13 provides some evidence of Granger-causality running from Eastern Cape to Southern Cape milk prices for a lag order of two to three months (which includes the optimal lag length) and across sample periods.

Table 12: Granger-causality tests between Western and Eastern Cape prices

Lag	January 2002 – December 2004		January 2002 – December 2005	
	Eastern → Western	Western → Eastern	Eastern → Western	Western → Eastern
1	**1.96** **(0.17)**	**0.18** **(0.68)**	**0.32** **(0.58)**	**2.58** **(0.12)**
2	4.92** (0.02)	0.13 (0.88)	1.52 (0.23)	1.11 (0.34)
3	3.19** (0.04)	1.66 (0.21)	1.38 (0.27)	2.44* (0.08)
4	2.94** (0.04)	0.96 (0.46)	1.03 (0.41)	1.58 (0.21)
5	1.91 (0.14)	0.78 (0.58)	0.64 (0.67)	1.29 (0.30)
6	1.98 (0.13)	0.74 (0.63)	1.27 (0.31)	0.90 (0.51)

Note: *** Reject at 1%, ** Reject at 5%, * Reject at 10%

Table 13: Granger-causality tests between Southern and Eastern Cape prices

Lag	January 2002 – December 2004		January 2002 – December 2005	
	Eastern → Southern	Southern → Eastern	Eastern → Southern	Southern → Eastern
1	0.41 (0.53)	0.66 (0.42)	0.64 (0.43)	0.09 (0.77)
2	3.12* (0.06)	0.37 (0.69)	3.77** (0.03)	0.07 (0.94)
3	**2.40*** **(0.10)**	**0.26** **(0.85)**	**2.57*** **(0.07)**	**0.05** **(0.98)**
4	1.92 (0.15)	0.38 (0.82)	1.73 (0.17)	0.04 (0.99)
5	1.72 (0.18)	0.65 (0.67)	1.60 (0.19)	0.07 (0.99)
6	1.50 (0.24)	0.47 (0.82)	1.47 (0.23)	0.30 (0.93)

Note: *** Reject at 1%, ** Reject at 5%, * Reject at 10%

Table 14 provides no evidence of Granger-causality between Southern Cape and Western Cape milk prices, with some significant results scattered across lag lengths yet without any systematic pattern to support a positive interpretation.

Table 14: Granger-causality tests between Western and Southern Cape prices

Lag	**January 2002 – December 2004**		**January 2002 – December 2005**	
	Southern → Western	**Western → Southern**	**Southern → Western**	**Western → Southern**
1	**0.63** **(0.43)**	**0.04** **(0.85)**	**1.56** **(0.22)**	**0.12** **(0.72)**
2	2.55* (0.10)	0.06 (0.94)	2.67* (0.08)	0.22 (0.80)
3	1.64 (0.21)	0.11 (0.96)	1.73 (0.18)	0.28 (0.84)
4	1.66 (0.20)	0.57 (0.69)	1.20 (0.33)	0.52 (0.72)
5	0.96 (0.47)	2.84** (0.05)	0.84 (0.53)	0.93 (0.47)
6	1.99 (0.13)	1.69 (0.19)	1.81 (0.14)	1.13 (0.37)

Note: *** Reject at 1%, ** Reject at 5%, * Reject at 10%

The Granger-causality tests complement – not substitute – the correlation analyses presented in the previous subsection. As noted earlier, the Granger-causality test is usually applied with the sole aim of confirming the *existence* of a short-run, dynamic relationship (ignoring the extent of the relationship) and also imposes a one-way logic which enjoys little support in the market definition literature. The test results above suggest a dynamic relationship between Eastern and Southern Cape prices, which is consistent with the strong correlation results obtained earlier. The evidence of a strong contemporaneous correlation between Eastern and Western Cape prices also finds some support from the Granger-causality test (albeit much weaker). There are no convincing results for a relationship between Southern and Western Cape prices, which might undermine the suggestion that the Western Cape formed a single market with the other two regions.

As noted in the discussion of the different tests, the above analyses are based on short-run information only. Relying on short-run information only implies a specific interpretation of the HM test for market definition. Even if one is only interested in short-run relationships, econometric research shows that the exclusion of long-run information from short-run models creates misspecification error. Consequently, the following subsections consider econometric tests dealing with long-run information.

3.4.6 Price-ratio stationarity tests

As noted earlier, unit root tests may be carried out on the log of the ratio of milk prices in any two of these regions to establish whether the particular ratio contains a stochastic trend. Persistence in any of these ratios will indicate that the particular ratio does not revert to a long-run equilibrium value. If a long-run price constraint is the appropriate price standard for market singularity, such persistence would indicate that the two constituent regions do not form one market for competition policy purposes. In this case, the log price ratios are calculated for the following pairs of regions:

(i) Western Cape and Southern Cape (hereafter called the W:S ratio).
(ii) Western Cape and Eastern Cape (W:E ratio).
(iii) Southern Cape and Eastern Cape (S:E ratio).

As argued earlier, and as illustrated by the frequently conflicting results for different sample periods, the 2005 institutional changes are likely to have affected price relationships. The W:E and S:E ratios are plotted in Figure 14 and Figure 15 over the sample period including 2005. The graphs confirm the impact of the institutional change, showing the W:E ratio declining from mid-2004 and the S:E ratio growing throughout 2005. Although the behaviour of the W:E and S:E ratios are not necessarily due to the agreement, the identification problem requires a conservative approach. Consequently, the sample period of 2002-*2004* is used for all milk price ratios.

The graphical intuition regarding stationarity is not clear for these ratios. Without a deterministic long-run path, both series may have a fairly constant average and variance, indicative of stationary behaviour. However, the wave-like persistence in the S:E ratio may cause one to conclude otherwise.

Figure 14: Log price ratio between Western Cape and Eastern Cape, January 2002 – December 2005

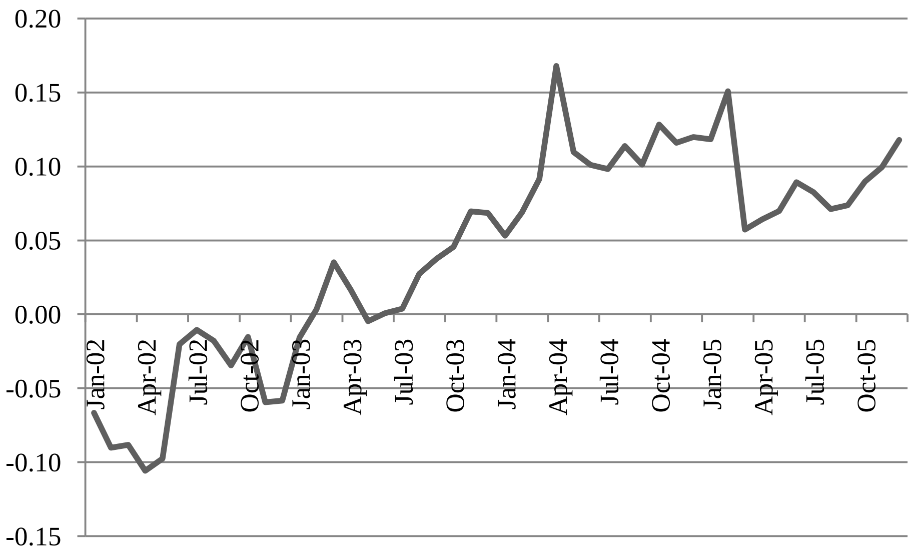


Figure 15: Log price ratio between Southern Cape and Eastern Cape, January 2002 – December 2005

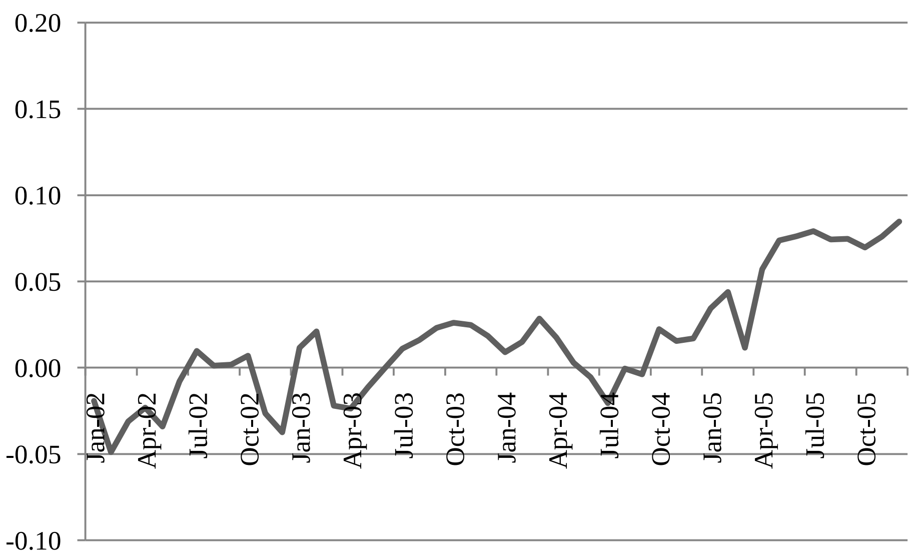
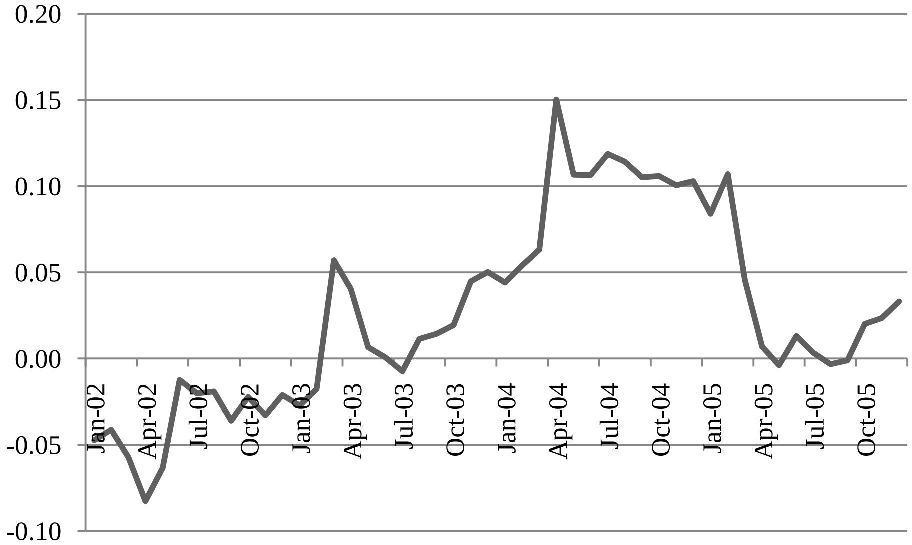


Figure 16 plots the W:S ratio. The W:S graph suggests non-stationarity, due to the strong persistence towards the end of the sample period. Figure 16 confirms the significant changes in 2005 during which Southern Cape prices were brought back in line with Western Cape prices, with a sharp decline in the W:S ratio during this period.

Figure 16: Log price ratio between Western Cape and Southern Cape, January 2002 – December 2005



The stationarity conclusions from Figure 14 to Figure 16 are ambiguous, and one can use formal unit root tests to evaluate the stationarity of milk price ratios. The Dickey-Fuller test results in Table 15 are strikingly different from the short-term conclusions and suggest non-stationarity for the W:S and S:E ratios and stationarity for the W:E ratio. However, as argued earlier, the DF results are likely to be misleading, given the small sample period. Consequently, the Ng-Perron tests have been applied and their results are reported in Table 16. The Ng-Perron tests confirm the results for the S:E ratio (non-stationary) and W:E ratio (stationary), but also indicate stationarity for the W:S ratio.

The stationary finding for the W:E ratio suggests a single market, which is the same conclusion drawn from the short-run findings. In other words, regardless of whether it is long- or short-run price co-movement that matters for market definition, evidence indicate that the Western and Eastern Cape milk regions constituted, at least until the point of institutional intervention, a single market. The non-stationary finding for the S:E ratio suggests separate markets, which is in contrast with the strong signals of market singularity provided by the short-run analyses. In a similar way, the finding of a stationary W:S ratio also contrasts with the earlier analyses, which did not indicate strong short-run relationships between the Western and Southern Cape (which is, possibly, the reason for the subsequent intervention).

Table 15: Augmented Dickey-Fuller (1979) tests on milk price ratios, January 2002 – December 2004

Lag	***t*-stat**	**p-value**
Western : Eastern		
1	-4.04**	0.02
2	-4.17**	0.01
3	-3.50*	0.06
4	-3.26*	0.09
5	-2.23	0.45
6	-1.76	0.70
Southern : Eastern		
1	-3.93**	0.02
2	-2.50	0.33
3	-2.96	0.16
4	-2.82	0.20
5	-1.84	0.66
6	-2.14	0.50
Western : Southern		
1	-3.67**	0.04
2	-3.07	0.13
3	-2.50	0.33
4	-2.18	0.48
5	-2.58	0.29
6	-2.33	0.41

Note: *** Reject at 1%, ** Reject at 5%, * Reject at 10%

Table 16: Ng-Perron (1996, 2001) tests on milk price ratios, January 2002 – December 2004

Lag	MZ_a	MZ_t	MSB	MP_T
Western : Eastern				
1	-25.37***	-3.55***	0.14***	3.65***
2	-91.18***	-6.75***	0.07***	1.02***
3	-217.99***	-10.44***	0.05***	0.43***
4	-1588.73***	-28.18***	0.02***	0.06***
5	-30.33***	-3.89***	0.14***	3.05***
6	-7.80	-1.96	0.25	11.72
Southern : Eastern				
1	-20.75	-3.22	0.16	4.39
2	-11.04	-2.35	0.21	8.25
3	-42.87	-4.63	0.11	2.13
4	-69.95	-5.91	0.08	1.30
5	-17.52**	-2.96**	0.17*	5.20**
6	-125.64***	-7.93***	0.06***	0.73***
Western : Southern				
1	-19.52**	-3.11**	0.16**	4.72**
2	-21.31**	-3.26**	0.15**	4.33**
3	-20.05**	-3.16**	0.16**	4.60**
4	-19.83**	-3.14**	0.16**	4.65**
5	-248.95***	-11.15***	0.04***	0.37***
6	-203.50***	-10.08***	0.05***	0.46***

Note: *** Reject at 1%, ** Reject at 5%, * Reject at 10%

The contrast between short-run and long-run results is one of the primary criticisms of econometric price tests raised by Coe and Krause (2008). However, as argued extensively, conflicting results do not necessarily indicate the empirical failure of a particular econometric test but may simply indicate that the different techniques ask different questions. Nevertheless, it is important not to take the results of the unit root tests as conclusive. Although the Ng-Perron tests have improved small-sample power, they have nonetheless been conducted on an extremely small sample, and more generally, make the potentially restrictive assumption that milk prices are related one-to-one (refer back to the methodology to see that tests on price ratios assume a cointegrating relationship of $(1; -1)$).

The preceding unit root tests rely on a sample period of January 2002 – December 2004 in order to deal with the problem of institutional change that directly affects price behaviour. However, given the very short sample period, the loss in power due to the loss of twelve data points is not negligible. Appendix D reports the unit root test results when conducted on the longer sample period including 2005. Restricting attention to the W:E and S:E ratios (given the impact of the institutional change in 2005 on the behaviour of the W:S ratio in that year), there is no evidence of long-run constraints using either the DF or Ng-Perron tests. The non-stationary result for the S:E ratio is consistent with the findings above, but the non-stationary result for the W:E ratio is the opposite.

A review of the pros and cons of including 2005 suggest that it is important to consider other econometric tests. Furthermore, even if one does find a long-run relationship, there is no guarantee that this long-run relationship is particularly important in monthly price changes – i.e. the speed with which a particular price adjusts to a long-run disequilibrium may be so slow as to be negligible and therefore of less importance in a market definition. Therefore, as argued earlier, it may be useful to pursue a technique that combines both short- and long-run information and also allows for a more general long-run relationship than the one-to-one relationship assumed here.

3.4.7 Bounds tests for level relationships and ARDL modelling

The bounds test and unit root test results on long-run relationships are not directly comparable and differ in three important ways. Firstly, the bounds test is based on an ARDL specification that is more congruent with the underlying data than the simple specification on which the unit root tests rely. Secondly, the bounds test is not limited to testing for a one-to-one long-run relationship. Thirdly, the bounds test relies on finite-sample critical values, as opposed to the asymptotic properties of the unit root tests.

Conditioning on both variables, one can fit bivariate ARDL models, first, to test for the existence of a long-run relationship and, second, to estimate the size of the long-run relationship where applicable. To enable valid inferences, each ARDL model is inspected for congruence with the data by running a batch of misspecification and diagnostic tests on the residuals (including tests for normality, heteroscedasticity, remaining autocorrelation and specification error). If the ARDL model passes these tests, the bounds test is performed.

The diagnostic tests identified a number of data outliers, which have been removed by including dummy variables. The inclusion of the outlier dummy variables significantly enhances the model fit (and also explains why the model results are likely to differ from the more simplistic stationarity tests or correlations, which are less concerned with deriving a congruent representation of the data). Table 17 lists the dummy variables employed in each case:

Table 17: List of dates for which dummy variables are included to remove data outliers

Relationship	**January 2002 – December 2004**	**January 2002 – December 2005**
Eastern → Western	April 2004	April 2004 March 2005 April 2005
Western → Eastern	November 2002 October 2004	November 2002 October 2004 March 2005
Southern → Eastern	November 2002 October 2004	November 2002 October 2004 March 2005
Eastern → Southern	January 2003	January 2003
Southern → Western	April 2004	April 2004 March 2005 April 2005
Western → Southern	January 2003	January 2003

Narayan (2005) calculated small-sample critical bounds for sample sizes $n = 35$ and $n = 45$. These are closest to the actual sample sizes of between 36 and 48. For comparison, the asymptotic values originally proposed by PSS are also reported. Table 18 reports the critical values:

Table 18: Critical *F*-values for bounds test

Significance level	$n = 35$		$n = 45$		**Asymptotic**	
	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
1%	7.87	8.96	7.74	8.65	6.84	7.84
5%	5.29	6.18	5.24	6.14	4.94	5.73
10%	4.22	5.05	4.23	5.02	4.04	4.78

Table 19 presents the results for the bounds test applied to the ARDL models with a lag order of six months. Results are similar for lower lag orders:

Table 19: ARDL bounds test results

Relationship	**2002-2004**	**2002-2005**
Eastern → Western	2.6027	1.1409
Western → Eastern	n.a.	n.a.
Southern → Eastern	3.3130	0.7364
Eastern → Southern	6.6261**	1.8329
Southern → Western	0.6791	1.3569
Western → Southern	0.6749	2.0022

** Significant at 95% confidence level

The results for the sample period including and excluding 2005 are consistent, with the exception being the Southern and Eastern Cape relationship. The bounds test results strongly suggest a long-run relationship between Eastern and Southern Cape milk prices in levels for the sample period excluding 2005, but not for the longer sample period. Other long-run relationships are not supported (even at a 10% significance level), but it is important to note that the F-statistics for the Western and Southern Cape equations rise significantly when including 2005, suggesting increased co-movement following the institutional change.

Given the significant result for Southern and Eastern Cape milk prices, it is possible to estimate the size of the long-run relationship. The model in Table 15 is not a parsimonious model and may contain, for example, irrelevant lagged variables that could have contaminated the long-run parameter estimates and could have led to a less robust model. Consequently, an automated GETS search algorithm is employed to reduce the GUM to a specific model. The algorithm chooses a number of starting points and, for each path, employs a step-wise reduction strategy to omit statistically insignificant variables provided information loss is limited (information loss is measured by change in the maximised log-likelihood value). The results of the multiple paths are then unified in a single model, on which the same step-wise reduction procedure is repeated until the model arrives at a single parsimonious model – known as the specific model.

Table 20 reports the ARDL results for the specific model, and Table 21 reports the relevant diagnostic tests.

Table 20: ARDL model of Eastern and Southern Cape milk prices (conditioning on Eastern Cape)

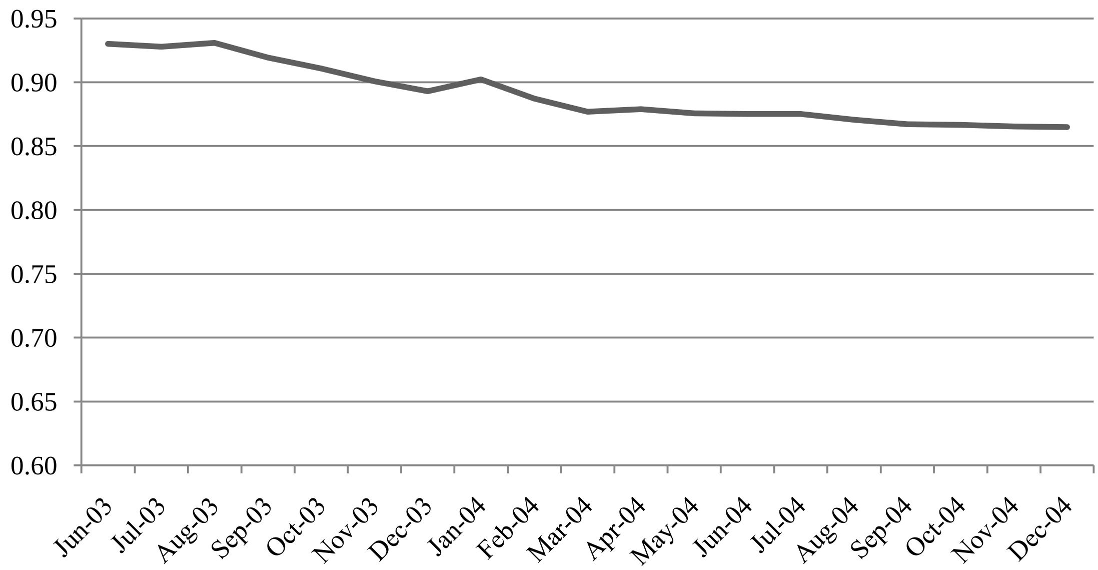
```
Modelling dsouth by OLS
The estimation sample is: 2002(7)-2004(12)

                   Coefficient  Std.Error  t-value  t-prob Part.R^2
dsouth_1              0.786044     0.1360     5.78  0.0000   0.6373
dsouth_3              0.187432    0.09722     1.93  0.0689   0.1636
dsouth_4             -0.210347    0.06405    -3.28  0.0039   0.3621
Constant              0.886349     0.2280     3.89  0.0010   0.4430
deast                 0.529314    0.06703     7.90  0.0000   0.7665
deast_1              -0.886894     0.1320    -6.72  0.0000   0.7039
deast_2              -0.374041    0.07749    -4.83  0.0001   0.5509
deast_3              -0.465180     0.1137    -4.09  0.0006   0.4684
east_1                 1.07399     0.1459     7.36  0.0000   0.7404
south_1               -1.24165     0.1601    -7.76  0.0000   0.7600
dum200301            0.0425731   0.008024     5.31  0.0000   0.5970

sigma               0.00657035  RSS             0.000820220119
R^2                   0.930319  F(10,19) =    25.37 [0.000]**
Adj.R^2               0.893644  log-likelihood        115.039
no. of observations         30  no. of parameters          11
mean(dsouth)        0.00218188  se(dsouth)          0.0201469
```

Table 21: Diagnostic tests on specific model

```
AR 1-3 test:       F(3,16)    =  0.89038 [0.4673]
ARCH 1-3 test:     F(3,24)    =   1.7305 [0.1875]
Normality test:    Chi^2(2)   =  0.29601 [0.8624]
Hetero test:       F(18,10)   =   1.1402 [0.4303]
RESET23 test:      F(2,17)    =  0.55318 [0.5851]
```

The results suggest a long-run correlation estimate of 0.86 (0.88 if based on the GUM) between Southern and Eastern Cape milk prices in the 2002 – 2004 sample period. This indicates that the levels relationship in the prices of the two regions is not only statistically significant, but also economically important. Given the short sample period, it is necessary to assess how robust the correlation estimate is to specific data points. Figure 17 reports the long-run correlation calculated from recursive estimates of the long-run parameters. The dates reported on the horizontal axis refer to the end dates for the recursive sample: the first sample ends at June 2003 (starting July 2002), and the sample is then incrementally enlarged by one month so that the final correlation is equal to the original sample used above. Figure 9 shows that the correlation estimate is remarkably stable over this short sample, declining slightly from around 0.93 to 0.86 as the sample grows, with the estimate remaining virtually unchanged when 2004 data points are added.

Figure 17: Long-run correlation between Southern and Eastern Cape based on recursive estimation, June 2003 – December 2004



3.4.8 Conclusion on the relevant market

The respondent employed some of the tests for price co-movement described here to argue for a single market. This chapter shows that the newer tests, including the ARDL bounds test, would have enhanced the market analysis by combining both short- and long-run information into a single model. The batch of tests presented here would have contributed to nuanced understanding of price relationships.

Table 22 summarises the results for the different price co-movement tests. The 2002 – 2005 sample period results for the Western and Southern Cape relationship are not discussed, given the direct impact of the institutional change on this relationship. Of course, it is likely that the institutional change in 2005 would have also affected the other two relationships, which is supported by the change in conclusions for the longer sample period.

Table 22 suggests the following conclusions. Firstly, if the focus is mostly on short-run price relationships, there appears to be a single market for the three regions. This follows from the significant result for both correlation and Granger-causality in the cases of the Southern and Eastern Cape and Western and Eastern Cape price relationships. Given these two relationships, and a chain-of-substitution argument, one may include all three regions in a single market. Secondly, if the focus is mostly on long-run relationships, Southern and Eastern Cape prices have a statistically and economically significant long-run relationship. Long-run relationships are also found between the Western and Southern Cape as well as between the Western and Eastern Cape prices, but these relationships do not appear to feature prominently in monthly price adjustment. The Southern and Eastern Cape can therefore be represented as a single market on the basis of long-run relationships.

Table 22: Summary of price test results

	January 2002 – December 2004			**January 2002 – December 2005**		
	W & S	**S & E**	**W & E**	**W & S**	**S & E**	**W & E**
Correlation	±0.5	±0.7	±0.5		±0.7	±0.5
Granger-causality	X	Short-run relationship	Short-run relationship		Short-run relationship	X
Price-ratio stationarity	Long-run relationship	Uncertain	Long-run relationship		X	X
Bounds test	X	Long-run relationship ±0.9	X		X	X

In sum, there is significant econometric evidence for a single market consisting of the three regions or, at least, a single market for the Southern and Eastern Cape. There is no evidence that the three regions should be considered separately. These conclusions are based on the 2002-2004 sample period. The results for the 2002-2005 sample period are different, for the reasons explained previously. However, given the nature of the institutional intervention, it is now more rather than less likely for the market to be integrated. Nevertheless, price-test results should be combined with other pieces of evidence. Within the framework presented earlier, price tests *combined* with other qualitative evidence on the SAMILCO agreement and on product flows support the identification of either a single market or two markets (one in the Western Cape and one in the Southern and Eastern Cape) for the south. Either of these conclusions implies a market share for the company under investigation of below 45% – weakening allegations of market power abuse against it.

3.5 Conclusion

This chapter considers the role of the law of one price in defining the relevant market for competition policy. In particular, it considers the following research question: are the various price-time-series tests for market definition consistent and can they be improved to reflect recent time-series developments? The chapter therefore investigates a range of conventional price tests used in market definition and introduces a number of alternative and improved price tests. The performance of conventional and new tests is then compared using data from a recent competition investigation in the South African milk industry. The main conclusions can be summarised as follows.

Firstly, some of the size and power problems faced by conventional tests are addressed by newer econometric procedures that are not yet widely applied. For example, based on the results in the empirical application, it is difficult to motivate the continued use (and therefore criticism) of the Dickey-Fuller test or contemporaneous correlation statistics when improved unit root tests or correlation analyses are available. These newer tests are easy to apply and quite similar in structure to existing price tests, which recommends their use in practice.

Secondly, as noted earlier, this chapter considers price tests for both short-run and long-run relationships, and each price test asks a specific question when evaluating a relationship. As shown in the empirical application, results can differ quite significantly. For example, compare the results for the unit root tests and ARDL models. Unit root tests on price ratios detect stationarity between Western and Southern as well as Western and Eastern price pairs, but are inconclusive for the Southern and Eastern pair. However, the ARDL model finds a long-run equilibrium only for the Southern and Eastern price pair. Nonetheless, the unit root tests and ARDL models ask quite different questions: the former focuses on the existence of a one-to-one relationship, while the latter focuses on whether a *general* long-run relationship *exists and is significant* in the month-to-month behaviour of the prices. Therefore, a decision concerning the validity of a particular price test for market definition depends on the particular price relationship question that the practitioner considers important for market definition. For example, practitioners must decide on the relevant time horizon for market definition when deciding which price tests are useful to apply: if the practitioner believes the SSNIP to refer to short-run behaviour, one could argue that he or she will prefer correlation analysis or Granger-causality tests. However, such a decision cannot be made a priori: ‘long-run’ adjustment can be as quick as six months or, alternatively, it could be relatively unimportant – implying a long run extending into a number of years. Therefore, it is advisable to run a batch of price tests that provide a rich perspective on price relationships, and then choose those tests which should be investigated further, based on the practitioner’s view of the appropriate hypothesis.

Thirdly, and related, the manner in which price tests are evaluated in the empirical application above illustrates that price tests ought to be used as exploratory tools for market definition, rather than as confirmatory tools. Price tests, as noted earlier, are based on a limited information set, and their results must be considered in the light of other descriptive or econometric evidence. Price tests are limited tools and cannot be used without reference to the context of the particular investigation. In the empirical application, for example, the institutional change at the start of 2005 suggests that price tests must be sensitive to the inclusion or exclusion of 2005 data. More generally, econometric tools in market definition, and competition policy more generally, should be understood as exploratory tools. It is a dangerous empirical approach to focus on finding a single encompassing test at the expense of the information offered by other types of tests. The noisiness of the data and the short length of the sample period are unlikely to support such an approach. This approach is consistent with the Popperian strategy adopted in most competition investigations, where evidence emerges in a piecemeal, evolutionary fashion. A variety of econometric tools assists in this discovery process. This chapter, together with Chapter 2, shows that where data challenges prevent more involved quantitative approaches, simpler tests such as cluster analysis and modern price tests can form part of such a discovery process.

A view of market definition as a discovery process suggests that market definition decisions are subject to significant uncertainty. The following chapter engages with this problem, by discussing the role of various types of uncertainty in market definition and by developing a conceptual tool (in the form a Bayesian decision rule) to assist practitioners in dealing with uncertainty when defining the relevant market.

Chapter 4

Market Definition as a Problem of Statistical Inference

In a competition case, the court or authority judges the welfare effects of a business practice or a proposed merger. The judgment by the court is an inference under uncertainty, and statistical decision theory can help to unpack the factors that influence such an inference. Cooper, Froeb, O'Brien and Vita (2005) use statistical decision theory to show that the rational judgment of a vertical restraint involves a careful weighing of pro- and anti-competitive evidence rather than a binary decision[15]. But the weighing of evidence is only one determinant of a court's judgment. Cooper et al. (2005) also show that judgment does not occur in a historical vacuum and that existing case law and economic research generate 'prior' probabilities that raise the standard of proof in any given case. In addition, the preferences of the decision maker (the loss function) also matter: the decision maker can assign different costs to the risk of disallowing a pro-competitive practice and of allowing an anti-competitive practice. Empirical evidence in a particular vertical restraints case can therefore be outweighed by case law, economic theory, and asymmetric loss functions – and such a judgment would still be considered rational.

Cooper et al. (2005) deals with the judgment of vertical restraints, but statistical decision theory can also be applied to market definition, which is subject to notable uncertainty. Statistical decision theory can help to unpack the roles of the weighing of evidence, prior probabilities and loss functions in market definition. Prior probabilities matter in this regard, as previous research and case law findings on substitutability raises the standard of proof in market definition. Loss functions may also be important, as the relevant market is defined by economists working for opposing parties, and they could weigh the evidence differently as a result of their side's perspective. When the weight of evidence is not strongly in favour of or against substitutability, prior probabilities and loss functions can determine whether the particular product is included in the relevant market. As shown below, an outcome dictated by prior probabilities or loss functions is consistent with a rational decision rule.

[15] In the United States, since the 1977 Sylvania ruling, complainants must link consumer harm to the vertical restraint; in the European Union, Article 81 and the 2010 vertical restraint guidelines emphasise an analysis of a particular vertical practice; section 5(1) of the South African Competition Act requires weighing of the pro- and anti-competitive effects of a vertical restraint.

This chapter studies market definition as a problem of statistical inference under uncertainty. Specifically, the chapter considers the following research question: given that market definition occurs under uncertainty, can one use a statistical decision rule to unpack the various factors important to a market definition decision?

To this end, the chapter is structured as follows. First, the chapter considers the need for an explicit treatment of uncertainty in market definition, focusing on conceptual ambiguity and model uncertainty. Second, the chapter introduces a Bayesian decision rule for market definition. Third, the chapter interprets the major elements of the Bayesian decision rule, emphasising the roles of prior probabilities, loss functions and weight of evidence. Fourth, the chapter discusses the problems of relying on a specific model for market definition. Finally, the chapter evaluates the extent to which market definition in a recently concluded South African merger case employed the decision rule principles developed in the chapter.

4.1 Uncertainty in market definition

The translation of economics into legal practice is a challenging task. Market definition is not exempt from this problem and is conducted under conditions of notable uncertainty, due to conceptual ambiguity and model uncertainty.

4.1.1 Conceptual ambiguity

Competition policy involves a combination of law and economics, with economics providing a conceptual framework for competition law (Neven, 2006). For example, the price elasticity of demand, monopoly and oligopoly are key concepts in competition policy, derived from economics (Baker and Bresnahan, 2008: 2). The application of rich economic concepts within the legal setting of competition policy poses challenges to practitioners and policymakers, as law requires codifiable concepts and unambiguous conditions in the interests of procedural efficiency and legal clarity. This need for legal clarity has led competition authorities in the US and EU to issue guidelines or directives to standardise economic analysis, by providing definitions of key economic concepts (see, for example, European Commission (2008), and United States Department of Justice and Federal Trade Commission (2010)).

Translating economic concepts into law is particularly challenging in the context of market definition. Substitutability is the central economic concept in market definition, but it is remarkably difficult to codify. As discussed in Chapter 1, US competition authorities implicitly define substitutability in merger investigations in terms of a thought experiment, the so-called hypothetical monopolist (HM) test (United States Department of Justice and Federal Trade Commission, 1992: 3):

> "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-maximising 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 HM test defines the substitutability of two products implicitly as the degree to which one firm is constrained by the product of the other firm in adjusting its price upward. This is a broad definition, but is intended to emphasise price constraints as the measure of substitutability. In its definition of product markets, the European Commission defines substitutability in terms of price characteristics, price relationships and use (European Commission, 1997):

> "A relevant product market comprises all those products and/or services which are regarded as interchangeable or substitutable by the consumer, by reason of the products' characteristics, their prices and their intended use."

In other jurisdictions, including the UK and developing countries such as South Africa, practitioners tend to rely on similarly broad definitions of substitutability (Theron, 2001; Davis and Garcés, 2010). The broad definitions adopted in the various jurisdictions reflect the complexity of the substitutability concept. However, broad definitions also reflect a trade-off: competition authorities weigh the need for legal clarity against the need for an informed analysis, which may require case-specific empirical analysis. This renders conceptual ambiguities, and the accompanying uncertainty, inevitable in market definition. Case law and economic research can identify market features that influence substitutability, which could partially address the problem of conceptual ambiguity. Despite these contributions, it remains difficult to codify exactly which features of a particular product should matter for substitutability, especially in the case of a differentiated product.

4.1.2 Model uncertainty

Apart from conceptual ambiguities, market definition also faces model uncertainty, i.e. uncertainty about the level of potential knowledge that an analyst can have about the model used to measure substitutability. As noted in previous chapters, 'model' in this context is broadly defined and covers fully specified econometric IO models but also a range of other tools used for assessing substitutability (including, for example, the price tests discussed in Chapter 3).

Model uncertainty is not limited to competition policy and is a common problem in applied economics. For example, model uncertainty is a "shared experience for applied econometrics" (Du Plessis, 2009: 428) and is closely associated with the debate surrounding the benefits and potential risks of data mining in econometrics (Hoover and Perez, 2000; White, 2000; Du Plessis, 2009). In the context of market definition, model uncertainty arises for two reasons: firstly, substitutability is unobservable and must be inferred from non-experimental data and, secondly, the model used for assessing substitutability relies on local rather than global information.

Substitutability is not directly observable and must be inferred from non-experimental data, which may be difficult to interpret. For example, the conditions assumed under the HM thought experiment are difficult to mimic in real-world markets: it is rare to find a natural experiment where the firm under investigation unilaterally raises its prices without other market factors also changing. Instead, analysts defining markets are forced to deal with the problem of inferring substitutability from observed data (Spanos, 2000; Baker and Bresnahan, 2008). But such inferences can be misleading, as observed data can be consistent with a number of possible interpretations. This problem of multiple interpretation pervades the social sciences and is known as the identification problem (Manski, 1995): it is difficult to 'identify' the correct model from the observed data. The identification problem forms the basis of Stigler and Sherwin's (1985) pessimistic label of the HM test as "completely nonoperational".

The identification problem in market definition is best explained by reference to a simple example. Suppose the analyst aims to study the price elasticity of demand for corn. It is usually not possible to directly estimate this demand elasticity, as the demand curve for corn is not observable. The analyst only observes the realised prices and quantities of corn traded, and it is not possible to identify a demand function from the price-quantity relationship without additional data. However, if additional data on supply-shift variables is available, the analyst can label particular quantity movements as demand movements. For example, suppose information on the weather conditions for a particular year is available. A match of annual weather conditions with the corresponding annual corn price and quantities would then help to isolate variations in quantities due to demand[16]. The weather variable is called an 'instrumental' variable. The identification problem therefore requires the analyst to search for a sufficient number of suitable instrumental variables.

The identification problem in market definition may be more complicated than suggested by the corn example, as market definition usually requires estimates on the cross-price elasticity of a variety of

[16] Geometrically, this happens because the weather variable splits the quantity-price data into two sets (each associated with a particular supply condition), which then helps to uncover the slope of the demand curve.

differentiated products. The identification problem is also not limited to cases where market definition relies on estimated econometric IO models for substitutability evidence: even when using predominantly anecdotal and qualitative evidence, the problem of inferring substitutability from the observable evidence remains.

The uncertainty created by the identification problem in market definition is exacerbated by the nature of the empirical tools employed in competition policy. Empirical tools tend to rely on local rather than global information, which restricts results to a subset of the information universe (Baker and Bresnahan, 2008). Practitioners often use qualitative evidence and simple quantitative tools with a limited information requirement (such as tests of price co-movement) when defining markets – especially given the time and data requirements of more rigorous tools, such as econometric IO models. The local information problem is not limited to cases without rigorous econometric evidence: even if presented, the court may not accept the *interpretation* of econometric evidence, or it may question the extent to which econometric evidence correlates with other pieces of evidence (Bishop and Walker, 1998).

A consequence of the local information problem is that it incentivises the non-disclosure of evidence: economists working for a particular party could focus on and elevate a particular aspect of substitutability favourable to their party's market definition. For example, in the South African radio merger case mentioned earlier, opposing parties disagreed in their interpretation of the demographic profile of radio listeners for the purposes of market definition. The intervening party grouped listeners into a larger group including all of the merging stations, whereas the merging parties differentiated more finely among listeners' demographic profiles to accentuate the differences between the merging stations.

Conceptual ambiguities and model uncertainty pose significant challenges to competition law investigations and, specifically, market definition. These uncertainties do not necessarily reflect poorly on the courts' ability to analyse evidence. Courts may perform reasonably well, given time, data and cognitive constraints. Nevertheless, a statistical decision framework that engages with the risks of incorrect inferences can assist in improving market definition decisions.

4.2 Market definition as a Bayesian decision rule

4.2.1 Using statistics for analytical purposes

Conceptual ambiguities and model uncertainty create significant risks of incorrect inference in market definition. Statistics can help to describe and manage these risks. For example, suppose the analyst relies on an econometric model to obtain price elasticity estimates for market definition purposes. The analyst can assess the risks of false positives and false negatives regarding the extent of substitutability, using the model output and statistical tests. But statistical tools can also offer analytical insights in settings with little quantitative data. Statistical decision theory represents one such a tool. For example, statistical decision theory played an important role in elucidating rational decision-making in the 1970s (Simon, 1979; Heiner, 1983). By emphasising the role of decision errors, statistics helped to explain the concept of bounded rationality. Statistical decision theory has also been applied extensively in competition law since the seminal paper by Easterbrook (1984). Easterbrook introduced the concept of error costs in an effort to incorporate the social costs of under- and over-enforcement in the development of optimal legal rules for competition policy (Manne and Wright, 2010). Since then, statistical decision theory has been used to derive optimal rules for a range of horisontal practices and mergers (see, for example, Hylton and Salinger (2001) and Beckner and Salop (1999)).

Statistical decision theory can also offer insights for market definition. A statistical decision rule for market definition can make critical decision factors explicit in terms of a rational framework. The development of a decision rule therefore requires an assumption of rationality.

4.2.2 Rationality

Market definition involves bounded rational decisions, in which the competition court or authority aims to use all available information optimally. In other words, the court weighs the available evidence and matches the evidence with current theory. Market definition as a bounded rational decision therefore requires neither perfect foresight nor full information – see Smith (2008) for a recent summary – as the court necessarily deals with limited evidence and limited knowledge. But how plausible is an assumption of bounded rationality for competition policy?

Competition policy decisions follow a process open to scrutiny and feedback, analogous to the process of scrutiny and feedback characterising scientific research. In science, the research presented by an individual must pass critical examination before it is acknowledged as scientific knowledge (Popper, 1963; Kuhn, 1970; Lakatos, 1977). Scrutiny and feedback take place via the so-called peer review mechanism, where peers of the researcher highlight alternative interpretations of data and ensure that the research is free from the subjective preferences of the individual researcher (Longino, 1998). Scientific research is therefore rational because of its social character, which encourages the critical assessment of research findings. It is from this perspective that one may argue competition court decisions to be rational: the critical feedback via the appeals process in competition policy introduces the kind of rationality obtained in a scientific process open to scrutiny and feedback. For example, in South African competition policy, the appeals process runs from the Competition Commission to the Competition Tribunal to the Competition Appeals Court (CAC), with further recourse to the Supreme Court and, ultimately, the Constitutional Court of South Africa (Competition Commission, 2000). This process involves a chain of decisions much less likely to fall short of rationality than any one legal decision.

The appeals process is directly relevant to market definition. The Kaya merger case discussed in Chapter 2 involved significant disagreement on the product market, and market definition was one of the central issues in the appeals that followed the Competition Tribunal's initial approval of the partial merger (Theron, 2010). In its first review, the CAC referred the matter back to the Tribunal, noting that the Tribunal failed to define a market in its first judgment. Therefore, the Tribunal in its second judgment explicitly delineated several markets, but still found in favour of the merging parties. Opponents of the merger still held that the market in the second judgment was not correctly defined and appealed again to the CAC. In its second review, the CAC concurred with the Tribunal's market definition and allowed the merger to proceed. This case offers one example of how the appeals process can influence competition policy decisions and, in particular, market definition.

Further support for an assumption of rationality comes from the requirement that any particular legal decision must be consistent with previous decisions. The consistency requirement is similar to the requirement for scientific research to be encompassing. A key criterion for a scientific contribution is that it incorporates previous results in addition to offering new insights, thus ensuring progress in science (Popper, 1963). For example, econometric models should encompass previous econometric results (Hendry, 1995). The encompassing criterion prevents ad hoc hypotheses without clear links to previous research from being accepted into the scientific body of knowledge. This criterion ensures the optimal use of available knowledge and thereby contributes to a rational outcome in scientific research. Analogous to scientific research, legal practice also promotes encompassing decisions: case law requires the consistency of a legal decision with previous decisions on the particular matter. The role of precedence in legal decisions therefore contributes towards a rational process in the adjudication of competition cases.

The rationality of court decisions has received extensive attention from legal scholars, as is witnessed in the academic debate between so-called rationalists and legal realists (Simon, 2004). Despite this disagreement, developing a decision rule for market definition may yet be useful. If the rationality assumption is disputed, it is important from a social welfare perspective to render competition court decisions more rational. The decision rule developed below can assist with this, by making critical decision factors explicit in terms of a rational framework.

4.2.3 Bayesian decision rule

A rational market definition exercise is based on evidence, both quantitative and qualitative, which can be denoted as e. As noted above, e includes all evidence that the court has access to and may not represent the entire set of available evidence: some evidence may not be reported by the parties involved and some evidence may not be admissible for legal reasons[17].

The process of market definition can be described as the application of a rational decision rule, which classifies a candidate product as either in or out of the relevant market based on e. Put differently, market definition is a binary classification problem – a candidate product is classified as inside or outside of the relevant market. Inclusion depends on the extent of substitutability between the candidate product and the product under investigation. When a candidate product is a substitute close enough to constrain the market power of the firm(s) under investigation, it is labelled as C, and is included in the relevant market. Alternatively, when it is not a close enough substitute to constrain market power it is excluded and labelled as F.

The decision to include or exclude a candidate product depends on the relative odds that a candidate product is a close substitute C. The relative odds are expressed mathematically as $\frac{P(C|e)}{P(F|e)}$, where:

$P(C|e)$ is the conditional probability of a product being a close substitute, given the evidence

$P(F|e)$ is the conditional probability of a product being a distant substitute, given the evidence

The Bayes (1783) rule states that, for two events A and B it can be shown that $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$. One can apply this rule to restate the conditional probabilities of being close and distant substitutes as follows:

$$P(C|e) = \frac{P(e|C)P(C)}{P(e)}$$

$$P(F|e) = \frac{P(e|F)P(F)}{P(e)}$$

The relative odds can then be rewritten as:

[17] Incomplete evidence may be less of a problem in adversarial competition jurisdictions such as the US and South Africa. Nevertheless, all competition investigations are subject to the local information problem, as discussed previously.

$$\frac{P(C|e)}{P(F|e)} = \frac{P(e|C)}{P(e|F)}\frac{P(C)}{P(F)}$$

The relative odds of the candidate product being included in the market are determined by the prior probability of being a close substitute, $\frac{P(C)}{P(F)}$, and the relative likelihood of the evidence being more consistent with what is expected of the evidence assuming the product is a close rather than a distant substitute, $\frac{P(e|C)}{P(e|F)}$.

These two terms explain the behaviour of different parties during market definition. Firstly, economists and legal practitioners will attempt to establish the prior[18] probabilities by relating extant IO research and quoting similar cases from the same or other jurisdictions, and the court will collate these arguments in forming its own prior beliefs about the merits of including a candidate product in the relevant market. Secondly, economists will attempt to uncover evidence for (or against) substitutability by means of quantitative or qualitative analyses and attempt to match the evidence with theoretical arguments concerning the size of the relevant market (Decker, 2009). The court will collate the different pieces of evidence and attempt to test whether the evidence is consistent with close substitutability or not.

The Bayesian exposition suggests that the courts may find the posterior probability (i.e. the relative odds) to be quite close to the prior probability of being included in the market when two conditions hold more or less simultaneously: inconclusive evidence *and* strong prior beliefs that the candidate product should be excluded. Put differently, when case evidence is inconclusive but previous research and case law strongly suggest that the product should be excluded, it is rational for the court to exclude it. This result implies that market definition is not necessarily an arbitrary exercise, as is often claimed. This issue receives attention later in the chapter.

What happens if there are no strong prior beliefs regarding substitutability *and* the evidence is inconclusive? Then the costs of a particular decision to include or exclude a product become important. In fact, these costs are always present, but become salient under conditions of high uncertainty, as the possibility of error is higher. Market definition can be interpreted as statistical inference, relying on the following hypotheses:

[18] 'Prior' is an unfortunate term in this context, as it may be (incorrectly) interpreted as denoting preferences. In this framework, preferences are captured by the 'loss function' (introduced later), while the prior probability refers to the scientific and legal record on the topic, i.e. information from the literature and case law developed 'prior' to the case (Salmon, 1998).

H_0: Product is C and should be included in the relevant market.

H_a: Product is F and should be excluded from the relevant market.

Market definition therefore faces type I error, excluding a substitute from the relevant market that would have constrained the market power of the firm(s) under investigation, and type II error, including a substitute in the relevant market that would not constrain the market power of the firm(s) under investigation. Let each error entail a specific loss: the loss associated with type I error is denoted L_1 and the loss associated with type II error is denoted L_2. Table 23 summarises the loss matrix:

Table 23: Loss matrix for market definition

		Product's actual substitutability	
		Close substitute	Distant substitute
Market definition decision	Exclude	L_1	0
	Include	0	L_2

If each loss is likely to occur with a certain probability, the expected losses from type I and type II errors can be obtained by scaling the loss function with the probability of the respective losses:

$$E(Loss_1|e) = L_1 P(C|e)$$

$$E(Loss_2|e) = L_2 P(F|e)$$

The expected loss functions allow us to derive an optimal rule for market definition. An optimal rule requires the competition authority to exclude a candidate product from the relevant market if:

$$E(Loss_2|e) > E(Loss_1|e)$$

Is the optimality condition the relevant one for market definition? One may argue that the court is only interested in minimising the probability of misclassification and is unbiased between the two classification errors. Alternative optimal rules, embodying different assumptions about the loss function, are easy to derive from the general rule, as the latter encompasses a number of special cases (Johnson and Wichern, 2002).

In applications of actual classification techniques to quantitative data, the analyst usually specifies the form of the probability distribution up-front. As is the case for other quantitative tools – including most of econometrics – the assumption is frequently that of normality, producing a linear classification rule. In the context of market definition, this would imply linear market boundaries (i.e. the space of substitutes is effectively 'chopped' up). However, the optimal rule for market definition developed in this chapter does not rely on such a normality assumption. The optimal rule therefore holds regardless of the shape of the market boundaries.

It is useful to restate the expected losses as follows:

$$E(Loss_1|e) = L_1 P(C|e) = \frac{L_1 P(e|C)P(C)}{P(e)}$$

$$E(Loss_2|e) = L_2 P(F|e) = \frac{L_2 P(e|F)P(F)}{P(e)}$$

Using these forms, the optimal rule can then be expressed as:

$$\frac{P(e|C)}{P(e|F)} < \left(\frac{L_2}{L_1}\right)\left(\frac{P(F)}{P(C)}\right)$$

The optimality condition suggests three conditions under which a product is excluded from the relevant market:

(i) $\frac{P(e|C)}{P(e|F)}$ is relatively small, i.e. the substitutability evidence is more likely to have been generated by a product not being a close substitute for the product(s) under investigation

(ii) $\frac{L_2}{L_1}$ is relatively large, i.e. the cost of including an irrelevant substitute in the market is high relative to the cost of excluding a relevant substitute

(iii) $\frac{P(F)}{P(C)}$ is relatively large, i.e. the decision maker has strong prior beliefs that a product should be excluded.

The Bayesian decision rule suggests that the case-specific evidence is not the only factor determining a competition authority's ruling on the extent of the relevant market: loss functions, prior probabilities, and the weighing of evidence all affect market definition decisions. The following section investigates each of these elements, highlighting what insights each offer for market definition.

4.3 Elements of the Bayesian decision rule and implications for market definition

Market definition is frequently contentious and the Bayesian decision rule can be a useful clarifying framework in practice. The rule highlights instances of less rigorous market definition decisions, i.e.

decisions failing to account appropriately for the available theory and case evidence, prior probabilities, or appropriate loss function. This decomposition renders court decisions more explicit by providing a simple taxonomy of the main elements of a rigorous market definition decision. The following subsections explore the different elements, considering how the loss function, prior probabilities and the weighing of evidence affect a rational market definition decision.

4.3.1 The loss function

The classical approach to inference does not account for the role of subjective preferences in shaping rational decisions (The Cambridge Dictionary of Philosophy, 1995). Under such an approach, it is difficult to judge whether a market definition decision is consistent with case law and literature or with case evidence, as it is difficult to distinguish between a rational market definition decision shaped by particular preferences and an incorrect market definition decision. A Bayesian approach to inference clarifies the role of preferences in a rigorous decision, by including a loss function in the decision rule.

The loss function describes the losses incurred from incorrect classification – either incorrect inclusion or incorrect exclusion from the relevant market. Arguably, the court may not have a bias in either direction, as it adopts the loss function implied in the competition law. As far as market definition is concerned, competition policy does not usually compel courts to define narrow or broad markets. Therefore, one could assume $L_1 = L_2$ in the optimality condition. This implies that, for an economist working for the court, the loss function does not necessarily feature in the market definition decision: the economist is led by the weight of evidence in a particular case and the evidence or theory from other cases and the literature (as embodied in the prior probabilities). However, the loss function may be an important driver of the proposed market definitions of opposing parties in a competition investigation. Opposing parties may arrive at relevant markets of quite different sizes even if substitutability evidence from neither is implausible. It may be that evidence supporting and evidence not supporting the inclusion of a specific product in the market are fairly balanced. However, opposing parties may deal differently with uncertainty about substitutability due to diverging loss functions. For one party, a loss-minimising approach would require excluding a specific product, while for the other, loss minimisation would require including it.

The loss function's role in a market definition decision serves to highlight the link between a specific market definition decision and the underlying competition policy regime. The divergence in loss functions between opposing parties may be particularly important under a form-based approach to competition policy. A form-based approach centres on establishing dominance and then proving that the conduct takes a particular prohibited form, rather than assessing the economic effects of the conduct. Under a form-based approach, much more hinges on the market definition, and the prosecuting authority will assign a heavier cost to defining a broad market, while the defendant will do the same for a narrow market. For example, where there is disagreement about substitutability, the prosecuting authority may favour excluding the substitute from the relevant market even if the evidence is slightly in favour of inclusion. Under an effects-based approach, dominance may be of a lesser concern (Arezzo, 2008). Arguments on the exact extent of the relevant market become less important and the loss function bias towards larger or smaller markets is reduced. Market definition therefore becomes less contentious because the loss functions of the parties converge (at least partially) on that of the court. Arguably, this convergence has two benefits.

Firstly, market definition and dominance can function more effectively as a first screen. It is frequently argued that market definition should be retained under an effects-based approach as it still helps to identify problematic cases. The argument for the retention of market definition is based on the assumption that, although the form-based approach has not been entirely successful, there is still evidence that structure yields certain conduct (Carlton, 2007). However, market share can be less useful as a screen for dominance when loss functions are asymmetric, leading the analyst to define a certain size market. The role of the loss function in market definition therefore highlights one of the familiar problems with using dominance as a first screen under a form-based approach: asymmetric loss functions may lead a prosecuting authority or another plaintiff to find that a firm holds significant market power and is capable of anti-competitive behaviour *because* the prosecuting authority may be inclined towards a narrow market definition.

Secondly, beyond ensuring a more effective first screen, the convergence of loss functions under an effects-based approach actually ensures that market definition reaches its true purpose, namely to identify and rank competitors. The market definition exercise should be seen as an information collection process, aimed at establishing and describing the space within which the firm's conduct or merger decision should be assessed. Converged loss functions ensure that prior probabilities and the weight of evidence dictate market definition.

4.3.2 Prior probabilities

Prior probabilities enjoy a prominent position in the philosophy of science, as scientists employ these probabilities in judging the plausibility of a new hypothesis (Salmon, 1998: 558-564). However, prior probabilities have not received much attention in the market definition literature. The Bayesian rule suggests that market definition, as any part of a competition investigation, does not occur in a historical vacuum. Strong prior probabilities concerning substitutability should raise the standard of proof in market definition exercises. The higher standard of proof does not necessarily affect the closer substitutes, but sets a higher bar for less close substitutes to be included, as the weight of case evidence for and against substitutability must be compared to prior probabilities. Furthermore, under conditions where the evidence is inconclusive, the Bayesian decision rule suggests that a product should be included in the relevant market if the prior probability of a close substitute is higher than the prior probability of a distant substitute.

The use of prior probabilities does not imply a subjective view of probabilities. Prior probabilities are to be established from a frequentist perspective, where 'probability' is measured by the frequency of a particular substitutability finding. For example, a high prior probability of being a close substitute implies that the larger part of the scientific record indicates close substitutability with the product under investigation. Of course, the use of these prior probabilities does not require exact numerical calculations. One may see the probabilities as intervals rather than as point values. While this may still leave room for divergence in views on prior probabilities, Salmon (1998: 564) notes that because of "washing out of the priors" these divergences need not affect the analysis negatively: as case evidence accumulates, the relative weight of prior probabilities in determining posterior probabilities declines.

The preceding discussion suggests that market definition could be enhanced if there were sufficient evidence on prior probabilities. The market definition component in most competition reports usually contains selected references to the literature, and it may be useful to develop administrative capacity to enhance the utilisation of extant research. Baker and Bresnahan (2008) suggest that economists can contribute to competition policy by developing a body of empirical evidence by industry type. The emphasis on prior beliefs suggests that competition regimes may benefit by developing the institutional capacity to collate the available evidence on the substitutability of particular products by industry type. The body of evidence should be updated frequently from new academic research and competition cases concluded elsewhere, and should be accessible to defendants. This need not involve interpretation in terms of market definition, but could focus mostly on conclusions regarding substitutability and the associated research methods.

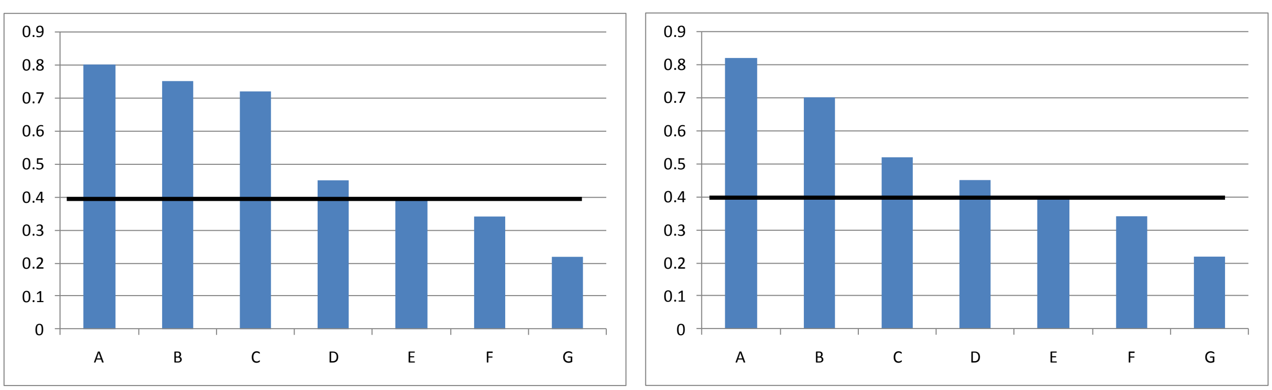
4.3.3 Weighing of probabilities

In the preceding discussions on the loss function and prior probabilities, the weighing of evidence received significant attention. In fact, the decision rule for market definition emphasises that a rigorous decision ultimately involves a weighing of probabilities for and against inclusion in the relevant market. Put differently, the strength of evidence must be explicitly considered during a market definition exercise. This is an important emphasis, as market definition is frequently applied in a binary sense, with substitutes being labelled either ‘in’ or ‘out’ of the market. However, some substitutes may be classified as ‘in’ on a much weaker basis than others. These substitutes may belong in the market, as they constrain market power, even though they represent a different level of substitutability relative to those products included on the basis of much stronger evidence.

The fact that the relevant market includes products with varying degrees of support (in terms of weight of probability) suggests that the rigorous drawing of market boundaries involves first the ranking of products and *then* the delineation of the market boundary. Such a perspective implies that the frequent calls for a less rigid approach to market definition need not involve loose market boundaries. A less rigid approach to market definition can be seen as one where there is a nuanced understanding of what lies *within* the explicit market boundaries. Such an interpretation balances the need for legal clarity and the need for minimum loss of information.

Figure 18 shows the information loss when defining the market on the basis of a comparison of the cross-price elasticity estimates with a pre-specified threshold level of, say, 0.4. On the left-hand side graph, three close substitutes and a distant substitute are included in the same market; on the right-hand side graph, the products are much more differentiated. An ‘in or out’ approach would treat these two scenarios as equal. The nuanced approach emphasised by the decision rule acknowledges that both markets include four products, but that the left-hand side market is more competitive (i.e. there are more constraints on the hypothetical monopolist) than the right-hand side.

Figure 18: Cross-price elasticity estimates for close and distant substitutes (left) and slowly decaying levels of substitution (right)



As argued earlier, a view of market definition as involving a weighing of probabilities does not necessarily involve quantitative calculations. The point is not that an optimal rule requires the decision maker to weigh mathematically a set of de-linked pieces of evidence. For example, the implication is not that the decision maker finds ‘eight’ pieces of evidence in favour of including the product in the relevant market and 'five' against. Decker (2009) discusses the way in which applied economists in competition cases weave evidence together. It is rarely the case that economists rely on a single table of cross-price elasticity estimates, which settles the market definition inquiry. Economists rely on a range of evidence weaved together to make the claim for (or against) substitutability. The substitutability evidence of any party is rarely outright false, and usually, it is the case that the evidence of one party is merely more compelling than another's – which implies weighting by the court. Furthermore, even if only econometric estimates are allowed as evidence for market definition, the weighing of probabilities *still* applies. Despite more extensive knowledge about substitutability and access to more sophisticated models, the application of knowledge to a specific case cannot be objective. Econometric models do not remove the need for judgment, given that they involve a large number of assumptions and rely on specific datasets. Competition courts will always have to weigh evidence regardless of the statistical sophistication of these models.

The argument that market definition involves the weighing of various pieces of evidence is relevant to South African competition policy, as it emphasises the need for strong economic evidence to underlie the determination of a particular market and assessment of market power. Consider, for example, the 2009-approved merger between South African wholesalers Masscash and Finro (Competition Tribunal, 2009). Although the case may appear to have focused less on the issue of a specific market delineation, a substantial portion of the case was devoted to the application of quantitative techniques measuring competitive constraints. In its assessment of these techniques, the Tribunal found that the apparently sophisticated statistical analyses did not hold up to closer econometric scrutiny or were contradicted by other qualitative evidence. The Tribunal's judgment suggests two lessons related to market definition. Firstly, model uncertainty is a common experience in applied econometrics and is not limited to market definition. The type of econometric procedures used in the case are frequently argued to obviate the need for formal market definition, yet these tools are also exposed to model uncertainty. Secondly, one should not engage one-dimensionally with market definition or any issue in competition policy. Any one econometric tool forms part of a broader kit of complementary tools and can never constitute the whole of the analysis. These insights are consistent with the argument that market definition involves a weighing of the probabilities of close and distant substitutability.

The weighing of probabilities during the market definition exercise suggests that it is sub-optimal to view market definition as a strict two-tiered process, where qualitative or crude quantitative evidence are treated as exploratory or indicative, while subsequent econometric work is seen as providing confirmation (see, for example, reports on methodology (Copenhagen Economics, 2003)). Instead, one may also view econometric evidence as one piece of evidence that may not necessarily weigh heavier than other pieces of evidence[19] (Bishop and Walker, 1998). This implies that market definition should not rely on a particular econometric or other quantitative technique. If one views the Bayesian decision rule as a statistical classification rule, multiple tools are preferable for market definition. The following section explores this issue.

4.4 The need for a variety of tools and evidence rather than a model-specific approach

The SSNIP thought experiment, a widely-used analytical tool for market definition, is frequently implemented via an econometric model (see Motta (2004: 125-134) and Davis and Garcés (2010: 204-227) for a summary of the approach). The SSNIP test is seen as requiring a formal estimate of elasticity for market definition, and advances in empirical IO modelling make it increasingly easy to obtain such elasticity estimates. The Bayesian decision rule presented earlier can also be interpreted along econometric modelling lines: the practitioner specifies an IO demand/supply model, estimates cross-price elasticities and then tests whether these elasticity estimates exceed certain thresholds, to decide whether a product should be included in the relevant market.

While econometric modelling provides useful evidence for market definition, it is potentially limiting to interpret the Bayesian decision rule as best implemented using a single econometric model. A key requirement for the use of such a model in defining markets is that it should enable not only retrospective analysis, but also forecasting. Market definition is inherently a forecasting exercise: for example, under the SSNIP test, market definition requires practitioners to predict the likely response of market players to a price hike implemented by a hypothetical monopolist. In econometric terms, the SSNIP test requires practitioners to generate conditional forecasts of the quantity effects of an exogenous change in price. But conditional forecasting opens up a range of empirical identification problems, including the need to establish whether price in any particular case can be viewed as strictly exogenous (Hendry, 1995). It is therefore remarkably difficult to utilise a single econometric model for forecasting in IO – see a recent discussion by Coate and Fischer (2011) on the evaluation of structural model forecasts in merger analysis.

Despite these problems, forecasting performance can be improved when a number of models are collated or when information from a variety of sources is combined (Granger and Jeon, 2004; Pesaran and Timmermann, 2005). The improvement follows because even the best performing forecasting model does

[19] This difference in approach is related to a broader debate on whether economics is a scientific discipline in the Popperian sense or a discipline employing the 'art' of economic story-telling using numbers and qualitative information (Leamer, 2009).

not necessarily encompass all other models so that a combination of models usually outperforms a single model. This suggests that market definition, as a forecasting exercise, is best conducted using multiple tools rather than a single econometric model, regardless of the complexity of such a model.

The classification literature sheds further light on the inherent problem facing the use of a single econometric model for market definition. Consider the ratio $\frac{P(e|C)}{P(e|F)}$, forming part of the decision rule for market definition. The court must judge and weigh these probabilities when defining the market. The judgment and weighing of probabilities rely on economic theory and case law. From a statistical classification perspective, one may say that the court classifies products as in or out of the market on the basis of a classification rule that is developed from economic theory and case law. However, classification rules are only as good as the data on which they are based. In cases where there is limited theory and case law dealing with the substitutability of the type of product under investigation, the classification rule for market definition may be biased. The statistical classification literature finds that unobserved variables lead to suboptimal classification rules (i.e. incorrect weighing of probabilities due to incorrect underlying theories) and incorrect classifications (Urbakh, 1971; Hastie et al., 2001). This outcome is likely in the case of many differentiated products, as only a portion of the characteristics driving substitutability is usually measured and/or considered when defining the relevant market.

These problems show the need for a range of data and evidence when classifying objects. The problem of limited information has led to the development of model averaging and boosting methods to improve on the decision rule suggested by a *specific* classification model (Hastie et al., 2001: 250). Therefore, similar to forecasting, a combination of classification models outperforms a single model. This lends further support to a market definition approach based on multiple tools.

More generally, the use of a single econometric model in defining markets builds on a social science philosophy that researchers can adequately model economic behaviour and devise improvements in markets[20]. Even if one supports this position, one must concede that knowledge is frequently outdated in the light of continuous change in the economy. Continuous change is important in market definition – for example, mergers can upset existing market structures and lead to changes in behaviour, which are not reflected in econometric models built on historical data. The on-going Wal-Mart/Massmart merger in South Africa is a useful example, with the entry of a large new supermarket player, Wal-Mart, likely to affect the behaviour of producers, other retailers and consumers. Price elasticity estimates of historical

[20] There appears to be a divide in economics between those favouring a rationalist position according to which social science has achieved adequate insights to motivate policy interventions that will effect social change, and those who are optimistic about the capacity of the decentralised process to solve social problems (Du Plessis, 2007).

consumer behaviour in the retail industry may therefore be less useful in defining markets. Continuous change in markets requires the use of a variety of sources and tools to collect sufficient and representative information. This links with the work on 'monocultural' competition policy by Budzinski (2008), who notes the limits of any one paradigm in explaining innovative activity and argues for diversity in economic paradigms, including those of Hayek (1946 [1984]) and Schumpeter (1934). The decision rule for market definition is therefore best implemented using a variety of tools and evidence.

It is useful to consider a recent competition case in which the competition authority engaged with the problem of uncertainty in market definition. Such a case study can show the extent to which market definition decisions, at least in South Africa, embody features of the statistical decision rule discussed above – even if the relevant authority does not explicitly use such a rule – and it will highlight areas for improvement.

4.5 Illustrative case: radio market definition in South Africa

The chapter illustrates the principles outlined above using the Primedia/Kaya radio merger case discussed in Chapter 2. As noted in that chapter, the case involved the partial acquisition of 24.9% of radio station Kaya FM (Kaya hereafter) by the media house Primedia and investment company Capricorn. Primedia already owned two other radio stations in the same province, Highveld Stereo (Highveld) and 702 Talk Radio (702). A significant part of the legal proceedings revolved around market definition, with the various parties in sharp disagreement regarding whether Kaya and Highveld (and 702 to a lesser extent) were in the same or separate relevant markets. The South African Competition Tribunal (Tribunal hereafter), in its final judgment, also acknowledged the importance of the market definition not only in facilitating concentration calculations, but also in "analysing incentives to co-ordinate" (Competition Tribunal, 2008: 21).

The Kaya case was discussed in Chapter 2 in the context of its relevance for cluster analysis and consumer characteristics. The case is also particularly relevant for the purposes of demonstrating how the Bayesian framework deals with uncertainty in market definition. As discussed below, the Tribunal explicitly recognised the uncertainty surrounding market definition for radio stations and, more important, attempted to account for this uncertainty in the market definition. The Tribunal sought to move away from a simple 'in' or 'out' approach to market definition and chose instead to rank substitutes of the merging radio stations using an 'inner' and 'outer' circle of competitors. The ranking of competitors and other features of the market definition exercise in the Kaya merger case approximate the approach suggested by the Bayesian decision rule derived earlier. The following subsections study the extent to which the various elements of the decision rule (prior probabilities, loss function, weighting of probabilities and variety of tools) featured in the Kaya market definition exercise.

4.5.1 Weighing of probabilities

In the preceding discussion of the elements of the Bayesian decision rule, the weighing of probabilities received significant attention: not all substitutes included in the relevant market enter with the same weight of evidence. For some substitutes, the evidence is much less certain than for others. A simplistic 'in or out' approach ignores the weight of probabilities, whereas the decision rule supports an approach that is explicit about these weights: as argued previously, this market definition approach starts with a ranking of substitutes, which is then followed by a formal delineation of the market boundaries. In this way, two objectives are achieved: on the one hand, a well-defined but general space is defined (useful for concentration calculations, etc) and, on the other hand, information about competition within the market boundaries is retained.

The market definition approach in the Kaya case illustrates the ranking approach. In fact, the Tribunal outlined the uncertainty associated with drawing clear market boundaries in the radio industry and agreed with the position of some parties that "it is far more meaningful to state propositions about relative relations between potential competitors than to make conclusions about absolute boundaries to markets" (Competition Tribunal, 2008: 15-16). The Tribunal therefore argued that "it is easier to ask whether A and B are more or less meaningful competitors in a market than, say, B and C than to ask which competitors must be regarded as in the market and which outside of it" (Competition Tribunal, 2008: 18).

The Tribunal applied the hypothetical monopolist test first to Kaya and then to Highveld to identify *and rank* competitors for each, as shown in Figure 19 (Kaya) and Figure 20 (Highveld). Specifically, the Tribunal ranked competitors by grouping competing radio stations for Kaya and Highveld into inner and outer circles of competitors, and placed those radio stations considered closest competitors for Kaya and Highveld in the inner-most circle (indicated with dashed lines in the graphs). The first outer circles (indicated with solid lines) contain substitutes also reasonably close, while the second outer circles contain less close substitutes that may still constrain price increases by the merging radio stations.

Figure 19: The market for Kaya as hypothetical monopolist

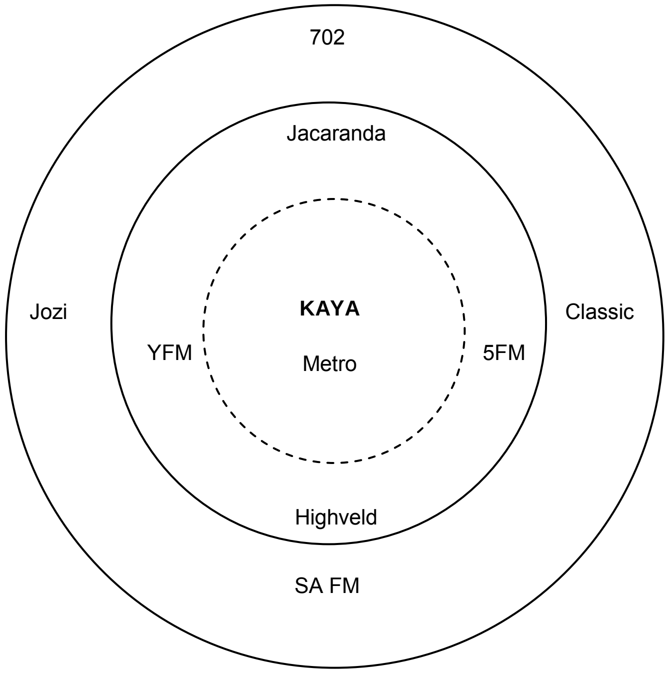
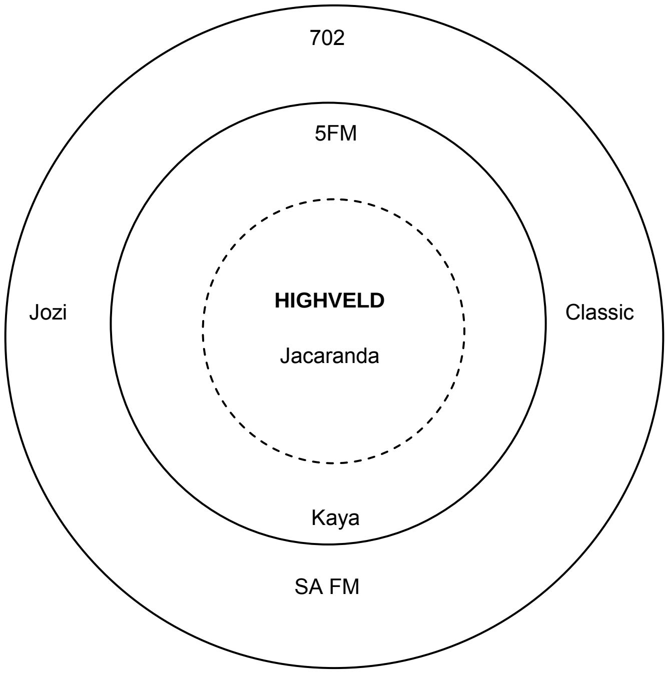


Figure 20: The market for Highveld as hypothetical monopolist



The outermost circles define the boundaries of the relevant market and can be used for concentration calculations. The various inner circles allow a nuanced assessment of competition: the Tribunal did not view all substitutes as equals and noted that "we must know something of who is in the market, *whether they have similar incentives to co-ordinate, and ... how much of the market might be party to the coordination*" (Competition Tribunal, 2008: 21, emphasis added). In fact, the Tribunal noted that there were conditions where only a part of the relevant market really mattered for Highveld: "We are not certain that the market extends only to the first circle, and there may be circumstances for some advertisers, as they are not a homogenous group either, that stations in the second circle are an adequate substitute, and this group of advertisers may be sufficiently large to deter a successful post merger price increase, assuming Primedia was in a position to control Kaya's pricing post merger" (Competition Tribunal, 2008: 20).

More generally, the Kaya case shows that uncertainty does not prevent rigorous market definition. Uncertainty does not require one to be vague about which substitutes are in the relevant market. It is best to define clear boundaries and then to be explicit about the ranking among the substitutes. This ranking approach is consistent with a statistical decision rule in which the probabilities of close and far substitution are weighed – implying that some products will be closer substitutes than others.

4.5.2 Loss functions

Apart from the weighing of probabilities, the loss function of the decision maker also affects rational decisions under a Bayesian decision rule. As discussed earlier, where concentration calculations significantly affect the outcome of a competition investigation, the loss functions, and hence market definitions, of opposing parties will tend to diverge strongly. This often occurs in mergers. For example, under South African competition law, mergers that do not significantly raise market shares and concentration (as measured by the Herfindahl-Hirshman Index) are permitted. Arguably, the Kaya merger case therefore offers a potentially good case study of how divergent loss functions affect market definition.

In its judgment, the Tribunal dismissed the more extreme market definitions put forward by the various parties. As discussed earlier, and consistent with a decision rule, this dismissal was based on the Tribunal's assessment of the evidence and, importantly, its own loss function (Competition Tribunal, 2008: 18): "At best the economists knocked holes into the more tenuous assumptions of their opponents, and so we can discount certain of the more border line theories of the relevant market …" The Tribunal, assigning similar losses to overly narrow and overly broad markets, therefore doubted proposals of either a single encompassing radio market or, alternatively, a collection of individual radio monopolies. In this sense, the Kaya case shows how the rational decision of an unbiased decision maker may differ from those of decision makers with specific biases.

Beyond confirming the existence of biases, the Kaya case offers further and potentially more important insights into the role of the loss function in market definition decisions. Most of the parties did not choose to pursue the extreme definitions. One could argue that it would have been almost expected that these extreme versions would not be accepted as the outcomes of serious analyses and that the issue was really how many substitutes to include before cutting off. As expected, the parties disagreed significantly. The Tribunal was forced to engage with the problem of judging the rankings suggested by the various parties and to come up with particular market boundaries: "we were still left with a middle ground [of economists' market definitions], where consensus could not be found and on whose assumptions, wildly conflicting notions of the extent of the concentration could be made" (Competition Tribunal, 2008: 18). Interestingly, while divergent loss functions would have influenced the parties' market definitions, the Tribunal did not appear to have considered the loss function a particularly important factor. The Tribunal noted specifically that the problem of radio market definition lay with the difficulty of weighing the evidence *rather* than with the partiality of the economists (Competition Tribunal, 2008: 18):

> "Radio markets are complex to analyse and although we had the benefit of the testimony and research of three economists in the course of the hearing in this matter, none of them emerged with a more probable version of the market than the others …. This is not a criticism of the efforts of these economists – indeed they all demonstrated great diligence in examining the data – but the nature of the industry".

In fact, the Tribunal admitted its own problems in interpreting the evidence (Competition Tribunal, 2008: 16): "the more data we received from the duelling teams of economists … the more the boundaries of the market receded into fog as opposed to clear lines".

This quote from the Tribunal is consistent with a Bayesian decision rule approach to market definition: the Tribunal argues that the loss functions of the various factors are unbiased and, therefore, that it is uncertainty in the data – formally, the weighing of probabilities – that is responsible for the different market definition conclusions. The quote also offers an important example of where uncertainty is not due to the decisions being arbitrary.

4.5.3 Prior probabilities

Apart from loss functions, prior probabilities also feature in the decision rule framework discussed earlier. Prior probabilities were not explicitly mentioned in the Kaya market definition decision, mostly due to the inconclusive nature of the scientific record on radio market definition[21]. For example, one market definition issue in the case concerned whether alternative media types should also be included in the relevant market. The Tribunal and most parties noted that other media (such as newspapers, the Internet or television) should not be included in the relevant market. The international literature does not necessarily offer clear guidelines in this regard. There was no formal empirical investigation and weighing of the case evidence in this regard either. Furthermore, while not publicised, the confidential reports of economists for opposing parties did consider the market definition suggested by previous cases in other jurisdictions. This would have contributed towards a situation where the prior probability ratio would have been close to one, and therefore immaterial in the assessment of the actual market.

4.5.4 Use of a variety of tools

Apart from the various elements of the decision rule, in the form of weighing probabilities, prior probabilities and loss functions, it has been argued that the decision rule suggested in this chapter supports the use of a variety of tools. In the Kaya case, econometric estimates of cross-price elasticity were not presented, yet market definition relied on a variety of other quantitative and qualitative tools. The Tribunal considered and weighed evidence from a range of quantitative analyses of demographics, but found that these offered inconclusive evidence on market definition: "[with] absent reliable econometric evidence and inconclusive evidence from the demographics of audience and advertiser profiles, the next best evidence we have of who competes in the market are the opinions of the stations themselves" (Competition Tribunal, 2008: 17). Subsequently, the Tribunal relied on qualitative evidence to form a view of the extent of the market. They avoided overreliance on any particular set of evidence, including the extensive demographic analyses of either side, and chose to weigh different pieces of evidence. This is consistent with the earlier argument that a decision rule for market definition is not model-specific but based on a range of quantitative and qualitative data.

The Tribunal's market definition approach in the Kaya case mirrors many of the key elements of a Bayesian decision rule approach to market definition. Notably, the weighing of the case evidence (via the various circles) and the explicit distinction between uncertainty and the loss function are key features of a rational market definition decision. The need for a variety of tools and, to a lesser extent, the importance of considering prior probabilities also received attention. While the Tribunal did not formally adopt a Bayesian rule, the Kaya case demonstrates that a Bayesian decision rule can be useful in defining markets under uncertainty. The Tribunal's decisions are generally consistent with the Bayesian rule's predictions,

[21] Arguably, prior probabilities were present in the negative: the Tribunal implicitly assigned a value of one to the prior probability ratio in commenting that the market is intrinsically difficult to analyse.

but a formal adoption of such a rule could have alerted the authorities to consider prior probabilities in greater detail.

4.6 Conclusions

The market definition decision is exposed to risks of incorrect inferences, due to conceptual ambiguity and model uncertainty. Conceptual ambiguity refers to the challenge of capturing the economic concept of substitutability in a form that is both legally codifiable and amenable to case-specific analysis. Model uncertainty concerns, firstly, the empirical identification problem of inferring substitutability evidence from non-experimental data and, secondly, the local information problem created by the inherently limited scope of our empirical tools.

Given these uncertainties, market definition faces significant risks of Type I and Type II errors. This leads to the research question: given that market definition occurs under uncertainty, can one use a statistical decision rule to unpack the various factors important to a market definition decision? The chapter suggests a role for a Bayesian decision rule, which can help to explain the factors that influence, or should influence, a rational market definition decision. The Bayesian decision rule highlights the role and interplay of the weighing of evidence, the loss function of the decision maker, and prior probabilities in determining the market definition decision.

Firstly, the decision rule underlines the importance of viewing market definition as a weighing of probabilities of close and distant substitutability. While the ideal types of ‘in the market’ and ‘out of the market’ are useful for constructing the actual boundaries, the rule shows that market definition first requires a ranking of substitution (by weighing probabilities) before actual boundaries are drawn. Not all products are included on the same strength of evidence.

Secondly, the decision rule shows that market definition need not be an arbitrary exercise. It is quite possible that different parties, using the same set of data, can disagree on the extent of the market: depending on the importance of the market definition to the case outcome, the loss functions of different parties may influence their market definition arguments.

Thirdly, the rule shows that the scientific and legal record on substitutability can enhance market definition decisions. The current literature on econometric tools for market definition makes little reference to the role of past cases and research, focusing mostly on improving the use of data in any specific case. The local and limited nature of evidence may be a barrier to proper market definition decisions, especially in differentiated product markets.

Apart from these conceptual contributions, the chapter makes a further methodological contribution regarding empirical techniques for market definition and, more broadly, for competition policy. Market definition is cast as both a forecasting exercise and a classification exercise. Building on the forecasting and statistical learning literatures, this chapter argues that optimal forecasting and classification are achieved when employing multiple tools. Therefore, the use of a single econometric model for market definition purposes is not desirable, and practitioners should attempt to use a broad toolkit rather than a single encompassing tool when defining markets.

The exposition of market definition from a statistical decision rule perspective is attractive to legal practitioners: portraying market definition as a weighing of probabilities is consistent with the legal standard of balancing probabilities, applicable in a number of competition jurisdictions, including South Africa. Given that the decision rule is consistent with legal principles, it could serve to improve future decisions by elucidating the factors that already play a role in decisions, thereby contributing to consistency of judgments. Apart from consistency and improved legal certainty, a decision rule approach also helps to understand the trade-offs made in arriving at a particular market definition, helping to address the charge that market definition is necessarily arbitrary.

Chapter 5

Conclusions

Economics studies the emergent order in an economy of production and exchange. This order is the result of innumerable transactions among millions of economic agents. Economists group these transactions using spatial concepts called markets. Each market refers to a specific collection of agents, sellers and buyers, transacting in a particular product or service. Competition policy is concerned with ensuring the competitive functioning of markets, by preventing large firms from artificially extending or exploiting their influence in markets. Therefore, as a first step, a competition investigation involves a careful definition of the extent of the relevant market.

Chapter 1 reviews the market concept in competition law investigations, defining it as a common space that includes all competitors constraining the market power of the firm under investigation. This definition differs from market concepts in other subdisciplines of economics and in the business world. Given the particular definition of 'market' in competition policy, market definition exercises share common features in the US, EU and South Africa. The various jurisdictions emphasise different aspects of substitutability, although demand-side substitutability generally enjoys greater weight compared to supply-side substitutability.

This background on the market concept in competition policy naturally leads to the fundamental question confronting research on market definition, namely whether market definition is a necessary first step in modern competition law investigations. Market definition retains its importance in merger cases for two reasons. Firstly, while market definition was originally designed for the assessment of market power via market share calculations, it is often also informative about the predicted effects of mergers because of its insights into substitution patterns. Secondly, market definition facilitates an eclectic empirical strategy: instead of relying on a single IO model to directly predict effects, the process of defining a market coincides with the use of a variety of quantitative tools and also qualitative evidence. This provides a richer picture and is preferred by practitioners.

Market definition also retains its usefulness in abuse of dominance investigations, even under an effects-based approach to anti-competitive practices. An effects-based approach requires the analyst to link supposedly anti-competitive behaviour with market effects. Critics therefore argue that market definition is no longer essential, as the concern now lies with effect rather than with proving that a firm is dominant based on market share. The conventional defence against this critique is that market definition is still a useful first screen identifying problematic cases, thereby economising on the efforts of the courts. However, there is a far more important role for market definition. Market definition, as argued above, involves the identification and ranking of substitutes. Substitution patterns are central to the effects-based approach: anti-competitive behaviour requires the use of market power, which only exists in the absence of meaningful competitors or the threat of their entry. The market definition exercise therefore assists directly in assessing the feasibility and possible effects of an allegedly anti-competitive practice.

If market definition is relevant to modern competition investigations, it still faces significant empirical challenges. Market definition relies on qualitative evidence and, increasingly, also on quantitative tools. Quantitative tools for market definition are relatively recent innovations in competition policy and became more sophisticated during the 1980s due to two developments. Firstly, the rise of cointegration and error-correction models motivated the use of more sophisticated time series tests of price co-movement. Secondly, at about the same time, the concept of residual demand laid the foundation for the increased use of fully-specified IO models for market definition and other purposes. While both promoted the use of quantitative tools, the two developments contributed to the emergence of quite distinct strands in the literature on market definition tools. The first strand is concerned with extending less data-intensive methods and implicitly promotes the use of a range of quantitative tools and evidence for market definition, while the second strand promotes a single encompassing tool for market definition in the form of a correctly specified econometric IO model. This dissertation is concerned with the first strand, given the notable practical and theoretical challenges facing IO models in practice.

Econometric IO models are attractive quantitative tools for market definition due to their price elasticity estimates, which are directly relevant to market definition. However, the application of IO models as exclusive tools for market definition face three important challenges. Firstly, as market definition forms only one part of market power assessment, it is not wise to spend a large amount of scarce resources on deriving exact estimates. In practice, competition authorities prefer a range of quantitative and qualitative evidence in market definition exercises, as the SSNIP test was never intended to be a technical statement on price elasticity but rather a description of the importance of evaluating competitive constraints. The problem with using formal elasticity estimates as the sole criterion for market definition is that one is implicitly comparing elasticity to some critical threshold. Secondly, the chapter highlights the practical constraints facing competition investigations: empirical IO models have significant data, time and capacity requirements that are frequently not met. Thirdly, and perhaps less important, the use of price elasticity estimates for market definition purposes faces the so-called cellophane fallacy problem in abuse of dominance cases, which may render the use of price elasticity in non-merger competition investigations theoretically problematic. Therefore, while empirical IO models can be useful, they are best used as part of a larger investigation based on a range of tools. A brief survey of the history of quantitative market definition tools in South Africa leads to three research questions:

(i) Research question 1: Can consumer characteristics assist in market definition and what tools can be used to compare the consumer characteristics of different products?

(ii) Research question 2: Are the various price-time-series tests for market definition consistent and can they be improved to reflect recent time-series developments?

(iii) Research question 3: Given that market definition occurs under uncertainty, can one use a statistical decision rule to unpack the various factors important to a market definition decision?

The first two research questions deal with the development of new and improved empirical tools for market definition: firstly, cluster analysis and consumer characteristics (Chapter 2) and, secondly, price co-movement tests (Chapter 3). The third question relates to a conceptual tool for market definition. Given the need for multiple tools, this dissertation engages with the problem of uncertainty in market definition. Specifically, it develops a Bayesian decision rule for rational market definition, highlighting the role of the loss function, prior probability and the weighing of evidence (Chapter 4). The following sections summarise the findings from Chapters 2 to 4.

5.1 Market definition using statistical learning techniques and consumer characteristics

Chapter 2 introduces the application of cluster analysis to data on consumer characteristics as a tool for market definition. Demand-side substitutability is the focus of market definition in competition law investigations, and this consumer focus implies that consumer data can be useful for market definition. In fact, direct forms of evidence on consumer behaviour, including consumer surveys and econometric models, are frequently employed during market definition. However, the practical constraints associated with these tools frequently prevent their wider application. In addition, the deeper theoretical assumptions of econometric IO models – including the view of competition as an outcome dictated by market structure – merit expanding the toolkit to include complementary tools that do not involve such assumptions.

The increased use of consumer characteristics in empirical IO models favours a tool that would allow a systematic analysis of consumer characteristics. IO models require data on individual consumer characteristics to improve elasticity estimates. However, even aggregated (such as market-level) data on consumer characteristics can already improve these estimates. This suggests that an analysis of average consumer characteristics for different products can be useful for market definition.

A study of consumer heterogeneity can be particularly useful in two types of cases: firstly, where systematic price data is not available (as is often the case in media markets where content is provided for free) and, secondly, where the product under investigation is closely targeted at a specific consumer group (so that consumer characteristics approximate product characteristics).

This background and motivation supports the use of a statistical tool for studying and comparing consumer characteristics. A comparison of the consumer characteristics of two products requires a study of similarity. If the consumer characteristics of a product can be represented mathematically by a vector, the similarity of the consumer characteristics of two products can be represented as the distance between two such vectors. The similarity of these consumer characteristics vectors is directly indicative of the extent of substitutability: the smaller the distance (the greater the similarity) between the two vectors, the greater the substitutability of the two underlying products. Therefore, if two products are in the same market, the two consumer characteristics vectors will tend to cluster together. Therefore, cluster analysis applied to consumer characteristics may be a useful additional tool for market definition.

A cluster is defined as a group of products, where the consumer characteristics are more similar to one another than they are to the consumer characteristics of other products outside of the group. Different analysts may identify different clusters, as the analyst determines the similarity concepts used during cluster analysis. Specifically, cluster analysis requires two similarity concepts: a distance metric (which deals with the similarity between individual products) and a linkage method (which deals with the similarity between clusters).

While a variety of distance metrics are available for cluster analysis, statistical distance best deals with scaling problems and correlation among individual consumer characteristics. Linkage methods for cluster analysis should be evaluated in terms of the need for both closeness and compactness in clusters: closeness ensures that the analyst does not include competitors in the market when there are closer competitors outside of the market, while compactness ensures that the market does not include too many irrelevant competitors. The closeness criterion therefore aims to minimise Type I error (incorrectly excluding substitutes), while the compactness criterion aims to minimise Type II error (incorrectly including substitutes). While the average linkage balances these two criteria, there are conditions where the analyst may want to emphasise one over the other. As different linkage methods may produce quite different clusters, the application of multiple linkage methods is preferable as it allows a comparison of results.

The distance measure and linkage methods are important inputs into the algorithm for hierarchical cluster analysis. Hierarchical cluster analysis does not generate clusters as an output, but instead reports a dendrogram. A dendrogram is a hierarchical tree diagram consisting of nodes, where the lowest level of nodes are the observations (different competitors), and these are merged successively (depending on their similarity) to form higher-level nodes until the entire data set is grouped as a single node. Each node therefore represents a potential cluster.

Hierarchical cluster analysis has five advantages as a tool for market definition. The first and foremost is that the dendrogram ranks competitors: ranking allows for a nuanced assessment of substitution patterns consistent with a view of market definition as an exercise in studying substitutability rather than merely for drawing boundaries. Secondly, the dendrogram does not require an a priori decision on the number of markets. Thirdly, the dendrogram enforces consistency by preventing analysts from forming arbitrary clusters that do not include all relevant competitors. Fourthly, the dendrogram is agglomerative (starting with individual competitors and successively merging these into ever larger clusters), which is consistent with the SSNIP logic of starting with one firm and successively including more competitors. Finally, hierarchical cluster analysis requires no distributional assumptions and has low sample size requirements.

Despite these advantages, the competition analyst may still need to draw market boundaries – and therefore require clusters to be identified. Several optimality criteria are available for tree cutting, each dealing with the optimal height at which to halt the clustering algorithm, taking the nodes at that height as the number of clusters. The gap statistic is the preferred optimality criterion. The statistic is based on the principle that the optimal number of clusters for a dataset is that number of clusters for which the total within-cluster variance for the data is lower than the total within-cluster variance of uniformly distributed data grouped into the same number of clusters.

The dendrogram and clusters suggested by the gap statistic are based on sample data and the analyst must ensure that sampling error has not compromised the analysis. A technique that allows statistical bootstrapping for the dendrogram can be useful for statistical inference in this context, given that the true population distribution is not known. Approximately unbiased bootstrapping procedures also address some of the statistical size problems experienced when relying on conventional bootstrap probabilities.

It is useful to illustrate the application of hierarchical cluster analysis in the context of an actual market definition problem. The illustration presented is based on a recently concluded South African merger case involving the South African media company Primedia acquiring a stake in Kaya FM, a Gauteng radio station. Primedia already owned two Gauteng radio stations Highveld and 702, and the merger proceedings centred on whether Highveld, being a so-called 'must-have' station, and Kaya FM share the same product market.

The Kaya case is a useful one, given a zero monetary price for radio content and the absence of reliable price data for advertisers, as well as the targeted nature of radio. The focus falls on the product market, as the starting point is a broad geographic market that includes both local Gauteng radio stations and national stations with a Gauteng footprint (i.e. stations with a signal that reaches a significant proportion of Gauteng). In addition, radio is assumed separate from other media, although the merits of this a priori decision are reconsidered in Chapter 4.

Data on consumer characteristics is obtained from RAMS and AMPS, providing some preliminary graphical comparisons of the listener profiles of Highveld, Kaya and 702. Cluster analysis on this data is performed using the average, single and complete linkage methods. The average and single linkage methods suggest the same substitutes, although the single linkage results are statistically more robust. In contrast, the complete linkage method suggests a larger market. While reconciling the different market definitions is challenging, it is not a problem unique to cluster analysis as a market definition tool.

Two sets of corroborative evidence support the market definition. Firstly, switching proportions, which refer to the proportion of listeners of a particular radio station listening to each of the other radio stations, was found to be consistent with the market definition suggested by the cluster analysis. Secondly, differences in the licensed radio formats of Highveld and Kaya constrain supply-side substitutability arguments and support the market definition suggested by cluster analysis.

One may raise a number of potential criticisms of cluster analysis and consumer characteristics as tools for market definition, including their lack of price data. In general, the counter-argument is that cluster analysis cannot be a substitute for a more sophisticated, or indeed any other, market definition tool, but represents a complement (Chapter 4 returns to this issue). Further criticisms of cluster analysis include the fact that cluster analysis deals with average but not marginal consumers and its assumption of mutually exclusive markets.

5.2 Advances in price-time-series tests for market definition

Chapter 3 deals with price co-movement tests for market definition. Price tests continue to be very popular in competition investigations, but have been subject to a range of empirical and conceptual criticisms. It is therefore useful to consider the merits of these criticisms and to introduce new price tests to deal with some of the criticisms.

The rationale for using price tests for market definition is based on the LOOP, which argues that prices in the same market should be equal (strong form LOOP) or related (the more popular LOOP version). The LOOP literature recognises the conditional nature of these relationships, citing inter alia the role of transport and transaction costs, exchange rates, and product differentiation. Despite these qualifications, critics maintain that the LOOP, and price tests by implication, is fundamentally unfit for market definition. Two main criticisms, one of a conceptual nature and another of an empirical nature, are often levelled against price tests.

The conceptual criticism of price tests holds that these tests cannot assist in defining markets for competition policy, as they focus on whether prices are linked rather than whether market players have the capacity to influence price. However, such a criticism sets too high a standard for a single tool – price tests should always be read together with other relevant evidence. Furthermore, as market definition is concerned with the ranking of substitutes, newer price tests can assist in this ranking as they consider not only the existence but also the size of relationships.

The second criticism is an empirical criticism of price tests, which argues that price tests have low statistical power in discriminating close and less close substitutes for market definition purposes. In addition, it is also claimed that simple price correlation outperforms more sophisticated price tests. Before considering the advantages and challenges associated with specific price tests for market definition, it is useful to highlight three general aspects related to the empirical criticism: (1) price tests are better viewed as tools for gathering information rather than for testing specific hypotheses – a common motivation for employing price tests in market definition is the inability to formally model price elasticity; (2) the empirical critique of price tests is based on theoretical models with restrictive DGPs that do not reflect what is commonly found in practical market definition settings; and (3) newer price tests with better power and size properties are available but have not yet been incorporated into studies critical of price tests.

Given this background, a variety of conventional and new price tests are available for market definition purposes. Each test asks a different question, and it is important to distinguish between these questions, given the frequent criticism that different price tests contradict one another (and are therefore less useful for market definition purposes). Each price test offers an alternative perspective on price relationships and, therefore, market definition.

Price tests can be classified into two groups: one set of tests aimed at studying short-run relationships and another set aimed at long-run relationships, also known as equilibrium relationships (and where ‘long-run’ is determined by the speed of adjustment towards equilibrium following a shock).

Short-run price tests include correlation and Granger-causality. The use of correlation statistics is problematic, given the need for a suitable benchmark when defining markets (e.g. is 0.4 high enough?). Correlation analysis can be expanded to consider the entire correlogram, i.e. not only the contemporaneous correlation between price series, but also the leading and lagging correlations. Despite this innovation, the main challenge with correlation analysis is one of interpretation, as it is often difficult to summarise the correlation evidence across the entire lag structure.

The Granger-causality test offers a way of systematically summarising information across the lag structure. One series is said to Granger-cause another series if the forecast of the other can be improved by including information from the first series. Granger-causality tests for market definition were popular in the 1980s, but this chapter highlights three challenges facing these tests. First, Granger-causality tests are usually employed to establish the direction of association, and it is not clear whether a bidirectional finding between two price series should be interpreted as stronger evidence of a single market compared with a unidirectional finding. Second, Granger-causality tests focus on establishing the existence rather than the size of a relationship. Third, Granger-causality tests require stationary series – given that most price series are non-stationary, the tests are usually run on first-differenced data, which implies a singular focus on short-run relationships.

Depending on one's interpretation of the LOOP, one may prefer long-run above short-run analyses for market definition purposes. Furthermore, ignoring long-run relationships may also bias the short-run analysis. One way of articulating a long-run relationship is in the form of co-integration: if a linear combination of two non-stationary price series is stationary, the two series are said to constrain one another over time and are called co-integrated. An elementary form of cointegration would involve a one-to-one relationship between the two price series such that the ratio of the prices is stationary. This implies that a test for stationarity (or a unit root test) on the ratio of the two price series can be used for market definition purposes.

Unit root tests have been subject to significant criticisms, which include the possibility of bias due to common shocks, already stationarity series, and product differentiation. These criticisms are not necessarily unique to unit roots and link back to the earlier comment that these tools cannot be employed in a vacuum separate from other quantitative and qualitative evidence. It is important, however, to ensure that enough data points are available over which mean reversion could occur, in order to test for long-run relationships.

Unit root tests often employed in market definition suffer from size problems (Type I error) and low power (Type II error) when sample sizes are small. These conventional unit root tests are therefore best avoided, given the availability of newer unit root tests with better size and power properties.

While unit root tests on the ratio of two price series are technically easy to apply, they face two limits. They do not consider, firstly, the possibility of a long-run relationship that is not directly proportional and, secondly, the relative roles of the short- and long-run in price behaviour. This suggests a role for error-correction models in market definition. However, while cointegration analysis can be used to build error-correction models, cointegration tests are subject to pre-testing bias (as one first has to verify that price series are all non-stationary in order to perform the cointegration analysis). As an alternative, an ARDL bounds testing approach may be used, as it does not require all price series to be non-stationary. The technique is also attractive in the data- and time-constrained environment of competition investigations, as it involves a single-equation regression model. In addition, small-sample critical values instead of the often employed asymptotic values are available.

It is important to illustrate the performance of conventional and new price tests in the context of an actual market definition problem. The illustration is based on a recently concluded South African competition investigation, where dairy processing firms were accused of contravening competition laws by concluding exclusive agreements with dairy farmers. One of the dairy processors operates processing plants in three southern milk regions (the Western Cape, Southern Cape and Eastern Cape), and it was necessary to establish whether these regions constituted a single market, in order to assess its dominance. Anecdotal evidence on product flows – among the southern plants and to a northern plant – strongly supports a single geographic market. Price tests can then be used to further explore this single-market hypothesis.

As expected, the empirical results are varied. Depending on the relevant time horizon, the various price test results differ quite significantly – with implications for the market definition conclusion. If practitioners believe short-run relationships to be the relevant issue for market definition, they will prefer the correlation and Granger-causality results. For this particular case, these tests support a conclusion of a single market across the three regions. However, as argued earlier, the results from the two types of tests are not always mirror images. The strong correlation results between the Eastern and Southern Cape are supported by strong Granger-causality results, but the similarly strong correlation between Eastern and Western Cape prices finds only weak support from the Granger-causality tests.

In contrast, if long-run price constraints are considered a more appropriate standard for market definition than short-run relationships, the chapter reports evidence from the bounds test of a statistically and economically significant long-run relationship between Eastern and Southern Cape prices. Unit root test results also detect one-to-one long-run relationships for the other two pairs of price series, but these do not feature prominently in monthly adjustment behaviour (i.e. the long-run relationship for these two pairs is swamped by short-run effects). Therefore, the long-run price tests are also not mirror-images of one another.

In sum, it is not advisable to choose a time horizon a priori. Long-run adjustment could be as quick as six months or, alternatively, it could be relatively unimportant – implying a long-run extending into a number of years. Practitioners are therefore advised to run a batch of price tests that provide a rich perspective on price relationships. Such a strategy is consistent with one of the remarks at the start of the chapter, namely that price tests should be considered exploratory tools. The noisiness of the data in competition cases and the frequent short length of the sample period do not support an approach based on a single encompassing test. The milk market test results provide a good illustration of this point.

5.3 Market definition as a problem of statistical inference

An expanded set of tools with low data requirements – including cluster analysis and price tests – can assist competition practitioners in defining the market, as many competition investigations proceed under constrained conditions: the market definition exercise is subject to data, time and cognitive limits. These constraints generate uncertainty, which leaves market definition arbitrary according to a long-standing criticism.

Uncertainty in market definition can be related to conceptual ambiguity and model uncertainty. Law requires unambiguous conditions and codifiable concepts in the interests of procedural efficiency and legal clarity. The need for such legal clarity is salient in market definition, as evidenced by the various guidelines or directives issued by competition authorities internationally to provide guidance on the definition of substitutability – arguably the central economic concept in market definition. The myriad of factors relevant to substitutability lead to fairly broad definitions, as competition authorities seek to balance the need for greater legal clarity against the need for a case-specific analysis.

Beyond conceptual ambiguity, model uncertainty is a significant source of uncertainty in market definition. Economists do not perform natural experiments but infer theories from non-experimental data subject to various influences. As far as market definition is concerned, the aim is to obtain information from observable data about substitutability, which is unobservable (or, at least, difficult to codify as discussed above). Consequently, there may be a range of interpretations of the substitutability evidence employed during a market definition exercise, which create a so-called identification problem. Furthermore, tools for market definition rely on local information due to time and data constraints, which further raises model uncertainty – and the local information also applies to rigorous econometric IO models. Local information may also incentivise economists working for various parties to focus and elevate a particular aspect of substitutability evidence favourable to their party's case.

To deal with uncertainty, it may be helpful to treat market definition as a problem of statistical inference. This requires a view of market definition as a bounded-rational decision. Rationality is a plausible assumption for market definition, due to the social character of legal decisions: the critical feedback via the appeals process introduces scrutiny and feedback, which are essential elements of a rational scientific process. Further support for an assumption of rationality comes from legal precedence, which ensures that legal decisions account for previous market definition decisions: by promoting encompassing decisions, legal practice creates a requirement for consistency similar to that found in scientific research.

A Bayesian decision rule can be useful in unpacking the various factors that affect a rational decision. Such a rule for market definition classifies a candidate product as either in or out of the relevant market based on the loss function of the decision maker, the prior probability of the product being in the market, and the weight of case evidence regarding substitutability.

The loss function describes the costs assigned to two types of errors in market definition: incorrect inclusion and incorrect exclusion from the relevant market. While courts and their economists may be unbiased between the two errors, opposing parties may assign heavier costs to one of these errors. The divergence of loss functions among various parties is accentuated under a form-based competition regime, which centres on establishing dominance (via market share calculations) and then proving conduct to take a particular prohibited form – with less emphasis on assessing the economic effects of the conduct. By implication, the rise of an effects-based approach will minimise the role of the loss function in market definition decisions, ensuring that these are increasingly driven by actual evidence rather than by preference. This contrasts with the conventional criticism of market definition as arbitrary.

Prior probabilities (of being a close or distant substitute) refer to the probability of a product being a close substitute based on scientific record and case law. A high prior probability of being included in the market implies that the larger part of the scientific record and the case law supports the inclusion of the market – independent of the actual case evidence. The importance of substitutability in competition law investigations suggests a need for institutional capacity by competition authorities to collect and maintain an accessible knowledge base regarding the substitutability of products or services.

The decision rule emphasises that a market definition decision also involves discussing and weighing evidence for and against inclusion in the relevant market so that some substitutes are included on a much weaker basis than others. If products are included with varying degrees of support, it is more appropriate to view market definition as involving the ranking of substitutes and then the delineation of market boundaries. Such an approach contrasts with the common view of market definition as a binary exercise where substitutes are labelled either 'in' or 'out'. Note that the weighing of probabilities need not involve only quantitative calculations. Furthermore, even if only econometric estimates were allowable evidence for market definition, the weighing of evidence would still be required – as econometric models involve a large number of assumptions and rely on specific datasets. Instead, econometric evidence forms part of a larger story – so that it would also be sub-optimal to see market definition as two-tiered, where qualitative evidence is treated as exploratory or indicative, and subsequent econometric analysis as confirmation.

The Bayesian decision rule is best implemented using a multiplicity of tools. Market definition is, firstly, a forecasting exercise and, secondly, a classification exercise – and both forecasting and classification are best performed using multiple tools. The use of a large empirical toolkit for market definition can also deal with continuous change and the problem of outdated knowledge.

The Competition Tribunal's market definition decisions for the Kaya merger case offers a useful case study to illustrate the principles of a Bayesian decision rule for market definition. Firstly, the Kaya market definition followed a ranking approach, grouping competitors for Kaya and Highveld into inner and outer circles depending on the closeness of substitutability. The multiple circles approach shows that uncertainty is best dealt with by defining clear boundaries and then being explicit about the ranking among the substitutes in the relevant market. Secondly, the Tribunal employed an unbiased loss function, which led it to reject the more extreme market definitions offered by opposing parties. Interestingly, prior probabilities did not feature prominently, with both the Tribunal and opposing parties agreeing that other media should not be included in the relevant market. However, there is no clear guidance from the international literature in this regard. Thirdly, the case involved the use of a variety of tools for market definition purposes. While econometric estimates of cross-price elasticities were not presented, a variety of other quantitative and qualitative evidence (including the analysis of consumer characteristics discussed earlier) were considered. In summary, the Kaya market definition meets most of the criteria for a rational market definition, although a proper consideration of prior probabilities may have improved the decision further.

5.4 Future research on market definition

The two empirical tools and one conceptual tool developed in this dissertation all raise further research questions. Furthermore, the focus in this dissertation is on markets in competition policy, which raises the question of whether and how the market definition exercise should differ in other policy settings. The following sub-sections consider each of these issues.

5.4.1 Cluster analysis and price co-movement tests as tools for market definition

The two empirical tools for market definition – cluster analysis and price co-movement tests – developed in this dissertation suggest a number of future research areas. The cluster analysis discussed in Chapter 2 focuses on the classification of data nodes (or items) on the basis of various node attributes. This classification approach assumes the mutual exclusivity of clusters: its aim is to organise the data into a single dendrogram with no overlap among the clusters. This may be a restrictive assumption from a market definition perspective. For example, in the analysis of a merger between firms A and B, it is not uncommon, during the application of the HM test, to find A in B's market, even while B is not included in A's market. If found, competition practitioners do not usually engage with such asymmetry and will tend to place A and B in a single market. Nevertheless, if the emphasis in market definition is to shift away from merely identifying substitutes towards also understanding substitution patterns *within* the market, it would be useful to account for such an overlap. The problem of overlapping communities has received significant attention in statistical physics, with numerous applications in the biological sciences, social networks, and in research on Internet search algorithms (Ahn, Bagrow and Lehmann, 2010). Multiple community membership precludes a unique hierarchical structure, and there is now a rapidly expanding literature on dendrograms (and, hence, clusters) of links (between pairs or nodes) rather than on nodes themselves (see Porter, Onnela and Mucha (2009) and Evans and Lambiotte (2009)). Link-based clustering derives from the empirical observation that, while nodes may belong to various communities, a group of links usually belongs to a single community. Investigating the use of link dendrograms for market definition purposes may therefore represent a promising future research area.

The price co-movement tests discussed in Chapter 3 suggest a number of areas for further research, of which two are highlighted here. Firstly, panel-based price tests may be useful in competition cases where price data is available for a large number of regions or products: in these contexts, panel tests offer substantial gains in statistical power. Boshoff (2007) investigates the use of the Levin, Lin and Chu (2002) and Im, Pesaran and Shin (2003) panel unit root tests for market definition in South African milk markets (see Chapter 3 for a related discussion). However, these panel tests are limited by their particular structure; the tests do allow testing hypotheses about unit roots in individual items within the panel and assume cross-sectional independence. However, newer panel tests can overcome some of these constraints (Pesaran, 2007; Pesaran, Smith, Yamagata and Hvozdyk, 2009; Gengenbach, Palm and Urbain, 2010). Investigating these tools for market definition may therefore be a useful avenue for future research – as they could offer significant advances even in samples of moderate size. Panel-based price tests can be particularly useful for the ARDL bounds test approach, which is a single-equation technique that treats price in one region or for one product as strictly exogenous (Hendry, 1995). Strict exogeneity of price is required for the SSNIP thought experiment, and it would be useful to consider, for example, a panel model that could account for potential endogeneity problems.

Secondly, the use of price data to define markets and assess market power can be further explored using analysis of variance (ANOVA) and panel estimation techniques. For example, suppose the practitioner has access to time series data on average prices charged by a variety of stores across a number of regions. It is then possible to assess the relative importance of, for example, store membership and region in price formation using ANOVA or, for a more dynamic assessment, by estimating various fixed effects in a panel regression model. This is useful for geographic market definition purposes, especially where there is disagreement about the relative roles of national and local factors in price setting, such as during the 2007 grocery inquiry in the UK. Some parties, including Tesco (one of the UK's largest grocery chains), argued for a national market, whereas other parties (including the UK Competition Commission) ultimately argued for a local market (see, for example Tesco's submission on geographic market definition (Hausman, 2007)). Market definition involved a number of tools (including price co-movement tests and extensive analyses of customer drive-time from stores), but there was no formal attempt at an econometric comparison of the *relative* roles of local and national factors in price setting (United Kingdom Competition Commission, 2008: 69-82). In the 2009-2011 supermarket probe in South Africa, the author attempted an ANOVA analysis for one of the supermarket chains, but the probe was concluded in January 2011 without extensive legal scrutiny of the empirical evidence (Competition Commission, 2011). Therefore, ANOVA seems to be a promising research area for price-based tests for market definition – especially since the type of regression models involved are relatively simple and easy to explain to legal professionals.

Thirdly, the time series nature of the price tests in Chapter 3 also links back to an older question in industrial organisation and competition policy, namely the usefulness of representative agent models in competition policy. These models do not account for consumer heterogeneity and can be difficult to estimate in markets with a large number of products, but have been retained in other settings, mostly due to the relative ease of estimation (see the discussion on the relative merits of various demand modelling approaches for competition policy purposes in Davis and Garcés (2010: 436-501)). A recent contribution by Coloma (2010) utilises a representative agent model to estimate so-called 'critical elasticity' and, consequently, to both define the market and predict horizontal merger effects. It may be useful to extend this analysis to abuse of dominance inquiries, bearing in mind the cellophane fallacy problem highlighted in Chapter 3.

5.4.2 Market definition as statistical inference under uncertainty

Beyond the future research areas linked to the two empirical tools, the conceptual tool developed in this dissertation – the statistical decision rule in Chapter 4 – also opens avenues for further inquiry. The decision rule framework developed in Chapter 4 emphasises that market definition is an inference under uncertainty, highlighting the various types of uncertainty associated with market definition. One area of possible further research concerns statistical uncertainty related to point estimates. A prominent example, already mentioned in Chapter 4, is the failed merger application by Masscash and Finro in 2009 (Competition Tribunal, 2009). In this merger, consumer surveys were used to estimate diversion ratios, in order to infer unilateral merger effects. The South African Competition Tribunal questioned the statistical validity of the point estimates, given the small sample sizes following stratification of the data. Future research may therefore focus on identifying the factors relevant to the assessment of point estimates and on developing a taxonomy of uncertainty in market definition.

5.4.3 Other research issues raised by the radio case example

The Primedia/Kaya merger plays an important role in this dissertation, providing a case example for both Chapter 2 and Chapter 4. In a policy context, the Kaya merger has also received significant attention for its market definition implications, as witnessed by a 2011 dedicated panel discussion on the merger (Mandela Institute, 2011).

Beyond the potential research areas suggested by the various tools, the Kaya case itself highlights two issues that merit further investigation. Firstly, as mentioned in Chapter 2, the case involves a two-sided market. Previous work by Boshoff, Du Plessis and Theron (2007) considers the implications of two-sidedness for market definition in the contexts of radio (the Kaya radio merger) and travel agent services (the SAA/Nationwide case). Given recent advances in formally modelling two-sided markets (Filistrucchi, Klein and Michielsen, 2010), this is a potentially important area for further research. It is especially relevant to the South African context, in light of a number of ongoing South African competition investigations dealing with two-sided markets. Two of these involve allegations of exclusive dealing, one in the context of Comair/Kulula's agreement with Lanseria airport (Competition Tribunal, 2010) and another in the context of Computicket's use of exclusivity agreements with some of its clients. This suggests useful case material for studying whether and how control of two-sided platforms can be used anti-competitively.

Secondly, beyond two-sidedness, the Kaya merger case also opens some questions concerning how market definition should be conducting in new or changing markets. As mentioned in Chapter 2, the argument that future technological change could join radio with other media types in a single market was not seriously entertained during the investigation. Furthermore, the Competition Commission argued that, in the longer run, the emergence of a significant Black middle class will render Kaya a significant competitor of Highveld – even if the two stations do not currently target the same audience. These issues are relevant also to current competition investigations – as mentioned above, the Comair/Kulula case concerns an agreement concluded with a newly-established airport (Lanseria) and an important question is therefore the extent of future competition between this airport and the existing O.R. Tambo airport. Therefore, the questions of whether a new product is to be included in a separate market and when it becomes substitutable enough may be some of the areas deserving of further investigation.

5.4.4 Market definition beyond competition policy settings

Beyond the specific research avenues suggested by the various tools and cases, the dissertation also raises the broader research question of how market definition in competition policy accords with market definition in other policy contexts, including telecommunications policy (Briglauer, 2008). In the South African context, market definition in telecommunications regulation has received significant policy attention – as was evident in the recent mobile termination proceedings at ICASA (Theron and Van Eeden, 2011). Theron and Boshoff (2006) identify a set of fixed-line and mobile telecommunications markets for South Africa, which can assist in analysing market power and other regulatory concerns in this industry. The authors highlight the practical difficulties in studying substitutability in an industry experiencing continued technological change: at that stage, the expected convergence of different telecommunications technologies posed a particular challenge to traditional market definitions of separate voice and data markets. While these dynamic considerations may be salient in the telecommunications industry, they are not unique to regulatory policy and, as mentioned previously, competition policy cases also contend with technological change when defining markets (Geroski, 2003).

More important than technological change is the question of whether the market concept in competition policy is even applicable in other policy contexts. As discussed in Chapter 1, a market in competition policy is a space worth monopolising so that the market is defined with the purpose of assessing market power. Telecommunications and other regulatory policies seek to promote competition in an *ex ante* fashion and may emphasise factors other than demand-side substitutability, which is usually the focus in market definition in competition policy settings (Briglauer, 2008). Nevertheless, at least part of the market definition problem in policy settings outside of competition policy appears to follow from its form-based approach. For example, ICASA assigns the label ‘significant market power’ to a provider based mostly on market share, with little further consideration of the linkages between markets and other factors influencing market power (Theron and Van Eeden, 2011). The research in this dissertation suggests that the policy emphasis on market share calculations is misleading; perhaps the issue is less one of changing the market concept in regulatory contexts, but of adopting an approach that is sensitive to the true role of market definition – namely the study of substitutability.

Reference List

Ahn, Y.-Y., Bagrow, J.P. and Lehmann, S. 2010. Link Communities Reveal Multiscale Complexity in Networks. *Nature*, 466(7307): 761-764.

Anderberg, M.R. 1973. *Cluster Analysis for Applications*. New York: Academic Press.

Arezzo, E. 2008. Is There a Role for Market Definition and Dominance in an Effects-Based Approach?, in M.-O. Mackenrodt, B. C. Gallego and S. Enchelmaier (eds.). *Abuse of Dominant Position: New Interpretation, New Enforcement Mechanisms?* Berlin: Springer. 5: 21-54.

Argentesi, E. and Filistrucchi, L. 2007. Estimating Market Power in a Two-Sided Market: The Case of Newspapers. *Journal of Applied Econometrics*, 22(7): 1247-1266.

Argentesi, E. and Filistrucchi, L. 2010. Market Definition in a Two-Sided Market: The Case of Newspapers. Unpublished paper delivered at the Competition and Regulation European Summer School and Conference 2010. Chania, Crete, Greece.

Baker, J.B. 2007. Market Definition. *Antitrust Law Journal*, 74: 129-173.

Baker, J.B. and Bresnahan, T.F. 1985. The Gains from Merger or Collusion in Product Differentiated Industries. *Journal of Industrial Economics*, 33(4): 427-444.

Baker, J.B. and Bresnahan, T.F. 1988. Estimating the Residual Demand Curve Facing a Single Firm. *International Journal of Industrial Organization*, 6(3): 283-300.

Baker, J.B. and Bresnahan, T.F. 2008. Economic Evidence in Antitrust: Defining Markets and Measuring Market Power, in P. Buccirossi (eds.). *Handbook of Antitrust Economics*. Cambridge MA: MIT Press.

Baker, J.B. and Rubinfeld, D.L. 1999. Empirical Methods in Antitrust Litigation: Review and Critique. *American Law and Economics Review*, 1(1): 386-435.

Baumol, W.J., Panzar, J. and Willig, R. 1982. *Contestable Markets and the Theory of Industry Structure*. New York: Harcourt Brace Jovanovich.

Bayes, T. 1783. An Essay Towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society*, 53: 370-418.

Beckner, C.F. and Salop, S.C. 1999. Decision Theory and Antitrust Rules. *Antitrust Law Journal*, 67(1): 41-76.

Berry, S.T. 1994. Estimating Discrete-Choice Models of Product Differentiation. *Rand Journal of Economics*, 25(2): 242-262.

Berry, S.T., Levinsohn, J. and Pakes, A. 1995. Automobile Prices in Market Equilibrium. *Econometrica*, 63(July 1995): 841-890.

Berry, S.T., Levinsohn, J. and Pakes, A. 2004. Estimating Differentiated Product Demand Systems from a Combination of Micro and Macro Data: The New Car Model. *Journal of Political Economy*, 112(1): 68-105.

Bhargava, A. 1986. On the Theory of Testing for Unit Roots in Observed Time Series. *Review of Economic Studies*, 53(3): 369-384.

Bishop, S. and Walker, M. 1998. *The Economics of E.C. Competition Law: Concepts, Application and Measurement*. London: Sweet and Maxwell.

Bishop, S. and Walker, M. 2002. *The Economics of E.C. Competition Law: Concepts, Application and Measurement*. London: Sweet and Maxwell.

Bjornerstedt, J. and Verboven, F. 2009. Predicting the Price Effects of a Merger in the Swedish Market for Analgesics: A Simulation Analysis. Unpublished paper delivered at the Competition and Regulation European Summer School and Conference 2009. Chania, Crete, Greece.

Blair, R.D. and Kaserman, D.L. 2009. *Antitrust Economics*. Oxford: Oxford University Press.

Boshoff, W.H. 2007. Stationarity Tests in Geographic Market Definition: An Application to South African Milk Markets. *South African Journal of Economics*, 75(1): 1-14.

Boshoff, W.H., Du Plessis, S.A. and Theron, N.M. 2007. Two-Sided Market Theory and Its Implication for Market Definition: An Application to the S.A.A./Nationwide and Primedia Cases. Unpublished paper delivered at the First Annual Competition Commission, Competition Tribunal and Mandela Institute Conference on Competition Law, Economics and Policy in South Africa. University of the Witwatersrand, Johannesburg.

Breiman, L., Friedman, J., Olshen, R. and Stone, C. 1984. *Classification and Regression Trees*. Belmont: Wadsworth.

Bresnahan, T.F. 1981. Departures from Marginal-Cost Pricing in the American Automobile Industry: Estimates for 1977-1978. *Journal of Econometrics*, 17(2): 201-227.

Briglauer, W. 2008. Conceptual Problems with the Hypothetical Monopolist Test in Ex-Ante Regulation of Communications under the New Regulatory Framework. *Journal of Competition Law and Economics*, 4(2): 311-334.

Brooks, G.R. 1995. Defining Market Boundaries. *Strategic Management Journal*, 16(7): 535-549.

Budzinski, O. 2008. Monoculture Versus Diversity in Competition Economics. *Cambridge Journal of Economics*, 32(2): 295-324.

Carlton, D.W. 2007. Market Definition: Use and Abuse. *Competition Policy International*, 3(1): 1-27.

Clark, T.E. 2006. Disaggregate Evidence on the Persistence of Consumer Price Inflation. *Journal of Applied Econometrics*, 21(5): 563-587.

Coase, R. 1937. The Nature of the Firm. *Economica*, 4(16): 386-405.

Coate, M.B. and Fischer, J.H. 2008. A Practical Guide to the Hypothetical Monopolist Test for Market Definition. *Journal of Competition Law and Economics*, 4(4): 1031-1063.

Coate, M.B. and Fischer, J.H. 2011. *Why Can't We All Just Get Along: Structural Modeling and Natural Experiments in Merger Analysis*. *SSRN Archive*. [Online]. Available: http://ssrn.com/abstract=1853675.

Coe, P.J. and Krause, D. 2008. An Analysis of Price-Based Tests of Antitrust Market Delineation. *Journal of Competition Law and Economics*, 4(4): 983-1007.

Coffey, A.J. and Sanders, A.K. 2010. Defining a Product Market for Spanish-Language Broadcast Media: Lessons from United States V. Univision Communications, Inc. And Hispanic Broadcasting. *Communication Law and Policy*, 15(1): 55 - 89.

Cogley, T. and Sargent, T.J. 2001. Evolving Post-World War I I U.S. Inflation Dynamics. *NBER Macroeconomics Annual*, 16: 331-373.

Coloma, G. 2010. Market Delineation and Merger Simulation: A Proposed Methodology with an Application to the Argentine Biscuit Market. *Journal of Competition Law and Economics*, 7(1): 113-131.

Competition Commission 2000. Policy Brief: Introducing the Act and the Competition Authorities. *Competition News*. Pretoria, Competition Commission. 1: 16-19.

Competition Commission 2011. *Competition Commission Findings of the Supermarket Industry Probe, Media Release 27 January 2011*. Pretoria: Competition Commission.

Competition Tribunal 2008. Clover Industries Limited and Clover S.A. (Pty) Ltd Versus the Competition Commission, Parmalat (Pty) Ltd, Ladismith Cheese (Pty) Ltd, Woodlands Dairy (Pty) Ltd, Lancewood (Pty) Ltd, Nestlé (Pty) Ltd and Milkwood (Pty) Ltd *78/CAC/Jul08*, South African Competition Tribunal.

Competition Tribunal 2008. Primedia Ltd, Capricorn Capital Partners Pty Ltd, New Africa Investments Ltd V Competition Commission and African Media Entertainment Ltd. *39/AM/MAY06*, South African Competition Tribunal.

Competition Tribunal 2009. Masscash Holdings (Pty) Ltd and Finro Enterprises (Pty) Ltd T/a Finro Cash and Carry. *04/LM/Jan09*, South African Competition Tribunal.

Competition Tribunal 2010. 1time Airline (Pty) Ltd (Complainant), Lanseria International Airport (Pty) Ltd (Respondent 1) and Comair Limited (Respondent 2). *91/CR/Dec09*, South African Competition Tribunal.

Cooper, J.C., Froeb, L.M., O'Brien, D.P. and Vita, M.G. 2005. Vertical Antitrust Policy as a Problem of Inference. *International Journal of Industrial Organization*, 23(7): 639-664.

Copenhagen Economics 2003. The Internal Market and the Relevant Geographic Market: The Impact of the Completion of the Single Market Programme on the Definition of the Geographical Market. *Report to the European Commission*.

Copenhagen Economics 2003. *The Internal Market and the Relevant Geographic Market: The Impact of the Completion of the Single Market Programme on the Definition of the Geographical Market*: European Commission.

Corbett, C., Das Nair, R. and Roberts, S. 2010. Countervailing Power, Bargaining Power and Market Definition: A Reflection on Two Mergers. Unpublished paper delivered at the Fourth Annual Competition Commission, Competition Tribunal and Mandela Institute Conference on Competition Law, Economics and Policy in South Africa 2010, University of the Witwatersrand, Johannesburg.

Cournot, A. 1927. *Researches into the Mathematical Principles of the Theory of Wealth*. New York: Macmillan.

Davis, P. and Garcés, E. 2010. *Quantitative Techniques for Competition and Antitrust Analysis*. Princeton: Princeton University Press.

Decker, C. 2009. *Economics and the Enforcement of European Competition Law*. Cheltenham: Edward Elgar.

Dewenter, R. 2003. The Economics of Media Markets. *Discussion Paper 10*. Helmut Schmidt University, Faculty of Economics.

Dickey, D.A. and Fuller, W.A. 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74(366): 427-431.

Du Plessis, S.A. 2007. Two Optimistic Traditions in the Dismal Science: Rationalism and the 'Invisible Hand'. *Inaugural lecture delivered on 28 February 2007*. Stellenbosch, Stellenbosch University.

Du Plessis, S.A. 2009. The Miracle of the Septuagint and the Promise of Data Mining in Economics, in H. Kincaid and D. Ross (eds.). *The Oxford Handbook of Philosophy and Economics*. Oxford: Oxford University Press.

Du Plessis, S.A. and Gilbert, E.S. 2007. Mark-Ups and Competition – a Comparison of the Profitability of Listed South African and American Industrial Companies. Unpublished paper delivered at the Economic Research Southern Africa Workshop on Institutions and Growth. 4-5 February 2008, Cape Town.

Du Plessis, S.A. and Gilbert, E.S. 2008. The Structure-Conduct-Performance (S.C.P.) Paradigm and Its Application in South Africa: A Review of the Empirical Evidence and the Implications for Competition Policy. Unpublished paper delivered at the Second Annual Competition Commission, Competition Tribunal and Mandela Institute Conference on Competition Law, Economics and Policy in South Africa. University of the Witwatersrand, Johannesburg.

Easterbrook, F.H. 1984. The Limits of Antitrust. *Texas Law Review*, 63(1): 1-40.

Edwards, L. and Van de Winkel, T. 2005. The Market Disciplining Effects of Trade Liberalization and Regional Import Penetration on Manufacturing in South Africa. *Working Paper 1*. Trade and Industrial Policy Strategies.

Efron, B., Halloran, E. and Holmes, S. 1996. Bootstrap Confidence Levels for Phylogenetic Trees. *Proceedings of the National Academy of Sciences USA*, 93: 13, 429–13, 434.

Ekelund, R.B., Ford, G.S. and Jackson, J.D. 1999. Is Radio Advertising a Distinct Local Market? An Empirical Analysis. *Review of Industrial Organization*, 14(3): 239-256.

Elliott, G., Rothenberg, T.J. and Stock, J.H. 1996. Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64(4): 813-836.

Elzinga, K.G. and Hogarty, T.F. 1973. The Problem of Geographic Market Delineation in Antimerger Suits. *Antitrust Bulletin*, 18(1): 45-61.

Engle, R.F. and Granger, C.W.J. 1987. Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 55(2): 251-276.

Europe Economics 2002. *Market Definition in the Media Sector: Economic Issues*. Brussels: European Commission: Director-General Competition.

European Commission 1997. Commission Notice on the Definition of Relevant Market for the Purposes of Community Competition Law. *Official Journal C 372 of 09.12.1997*. Brussels: 5-13.

European Commission 2008. Guidance on the Commission's Enforcement Priorities in Applying Article 82 E.C. Treaty to Abusive Exclusionary Conduct by Dominant Undertakings. Brussels, Directorate-General Competition.

Evans, D.S. 2009. Two-Sided Market Definition. *Market definition in antitrust: theory and case studies*. American Bar Association Section of Antitrust Law.

Evans, T.S. and Lambiotte, R. 2009. Line Graphs, Link Partitions, and Overlapping Communities. *Physical Review E*, 80(1): 016105.

Everitt, B.S. 1993. *Cluster Analysis*. London: Edward Arnold.

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

Fedderke, J.W. and Naumann, D. 2009. An Analysis of Industry Concentration in South African Manufacturing, 1972-2001. *Applied Economics*, 43(22): 2919-2939.

Fedderke, J.W. and Simbanegavi, W. 2008. South African Manufacturing Industry Structure and Its Implications for Competition Policy. *Economic Research Southern Africa Working Paper 111*. Economic Research Southern Africa.

Fedderke, J.W. and Szalontai, G. 2009. Industry Concentration in South African Manufacturing Industry: Trends and Consequences, 1972-96. *Economic Modelling*, 26(1): 241-250.

Felsenstein, J. 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution*, 39: 783–791.

Filistrucchi, L. 2008. *A S.S.N.I.P. Test for Two-Sided Markets: The Case of Media*. *SSRN eLibrary*. [Online]. Available: http://ssrn.com/paper=1287442.

Filistrucchi, L., Klein, T.J. and Michielsen, T. 2010. Merger Simulation in a Two-Sided Market: The Case of the Dutch Daily Newspapers. *Working Paper No 10-15*. NET Institute.

Fisher, F.M. 2007. Economic Analysis And "Bright-Line" Tests. *Journal of Competition Law and Economics*, 4(1): 129-153.

Fok, D., Franses, P.H. and Paap, R. 2006. Performance of Seasonal Adjustment Procedures: Simulation and Empirical Results, in T. C. Mills and K. Patterson (eds.). *Palgrave Handbook of Econometrics*. Hampshire: Palgrave Macmillan: 1035-1055.

Forni, M. 2004. Using Stationarity Tests in Antitrust Market Definition. *American Law and Economics Review*, 6(2): 441-464.

Fourie, F.C.v.N. 1996. Industrial Concentration Levels and Trends in South Africa: Completing the Picture. *South African Journal of Economics*, 62(2): 123-143.

Fourie, F.C.v.N. and Smith, A. 1993. Concentration, Tariff Protection and Industrial Performance in South Africa. *South African Journal of Economics*, 61(3): 116-133.

Fourie, F.C.v.N. and Smith, A. 1998. The Concentration-Profits Stalemate I: Causality and Interpretation Problems. *South African Journal of Economics*, 64(4): 558-583.

Fourie, F.C.v.N. and Smith, A. 1999. The Concentration-Profits Stalemate I I: Ideological Rifts and Methodological Stress. *South African Journal of Economics*, 67(1): 65-94.

Frankel, J., Smit, B. and Sturzenegger, F. 2008. South Africa: Macroeconomic Challenges after a Decade of Success. *Economics of Transition*, 16(4): 639-677.

Froeb, L.M. and Werden, G.J. 1991. Residual Demand Estimation for Market Delineation: Complications and Limitations. *Review of Industrial Organization*, 6(1991): 33-48.

Genesove, D. 2004. Comment on Forni's 'Using Stationarity Tests in Antitrust Market Definition'. *American Law and Economics Review*, 6(2): 476-478.

Gengenbach, C., Palm, F.C. and Urbain, J.-P. 2010. Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling. *Econometric Reviews*, 29(2): 111-145.

Geroski, P.A. 1998. Thinking Creatively About Markets. *International Journal of Industrial Organization*, 16(6): 677-695.

Geroski, P.A. 1998. Thinking Creatively About Markets. *International Journal of Industrial Organization*, 16: 677-695.

Geroski, P.A. 2003. *The Evolution of New Markets*. Oxford: Oxford University Press.

Gordon, A.D. 1999. *Classification*. London: Chapman & Hall / CRC Press.

Gorman, T. 1980. A Possible Procedure for Analysing Quality Differentials in the Egg Market. *Review of Economic Studies*, 47(5): 843-856.

Granger, C.W.J. 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 37: 424-438.

Granger, C.W.J. and Jeon, Y. 2004. Thick Modeling. *Economic Modelling*, 21(2): 323-343.

Granger, C.W.J. and Newbold, P. 1974. Spurious Regression in Econometrics. *Journal of Econometrics*, 2(2): 111-120.

Gual, J., Hellwig, M.F., Perrot, A., Polo, M., Rey, P., Schmidt, K.M. and Stenbacka, R. 2005. An Economic Approach to Article 82: Report to the European Commission. Brussels, European Advisory Group on Competition Policy (EAGCP).

Haldrup, N. 2003. Empirical Analysis of Price Data in the Delineation of the Relevant Geographical Market in Competition Analysis. *Working Paper 2003-09*. University of Aarhus.

Haldrup, N. and Jansson, M. 2006. Improving Size and Power in Unit Root Testing, in T. C. Mills and K. Patterson (eds.). *Palgrave Handbook of Econometrics Volume I*. Hampshire: Palgrave Macmillan: 252-277.

Haldrup, N., Montanés, A. and Sanso, A. 2005. Measurement Errors and Outliers in Seasonal Unit Root Testing. *Journal of Econometrics*, 127(1): 103-128.

Hastie, T., Tibshirani, R. and Friedman, J. 2001. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. New York: Springer-Verlag.

Hausman, J.A. 2007. *An Assessment of Geographic Market Definition in the U.K. Groceries Market*. *Tesco submission on geographic market definition in U.K. grocery market inquiry*. [Online]. Available: http://www.competition-commission.org.uk/inquiries/ref2006/grocery/pdf/main_party_submissions_tesco_geographic_market_definition.pdf.

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

Heiner, R.A. 1983. The Origin of Predictable Behaviour. *The American Economic Review*, 73(4): 560-595.

Hendry, D.F. 1995. *Dynamic Econometrics*. Oxford: Oxford University Press.

Holden, M. 2007. Market Relevance in a Small Open Distant Economy. Unpublished paper delivered at the Economic Society of South Africa Biennial Conference. Johannesburg.

Hoover, K.D. and Perez, S.J. 2000. Three Attitudes to Data Mining. *Journal of Economic Methodology*, 7(2): 195-210.

Horowitz, I. 1981. Market Definition in Antitrust Analysis: A Regression-Based Approach. *Southern Economic Journal*, 48(July 1981): 1-16.

Hosken, D. and Taylor, C.T. 2004. Discussion of 'Using Stationarity Tests in Antitrust Market Definition'. *American Law and Economics Review*, 6(2): 465-475.

Hunter, N.A. 2008. The Market Definition Paradigm. *Department of Economics*. Santa Barbara, University of California. Ph.D.: 118.

Hylton, K.N. and Salinger, M. 2001. Tying Law and Policy: A Decision Theoretic Approach. *Antitrust Law Journal*, 69(2): 469-526.

Im, K.S., Pesaran, M.H. and Shin, Y. 2003. Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics*, 115(July 2003): 53-74.

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.

Ivaldi, M. and Verboven, F. 2005. Quantifying the Effects from Horizontal Mergers in European Competition Policy. *International Journal of Industrial Organization*, 23(9-10): 669-691.

Johnson, R.A. and Wichern, D.W. 2002. *Applied Multivariate Statistical Analysis*. Upper Saddle River: Prentice-Hall.

Johnson, S.C. 1967. Hierarchical Clustering Schemes. *Psychometrika*, 32: 241-254.

Katz, M.L. and Shapiro, C. 2003. Critical Loss. Let's Tell the Whole Story. *Antitrust Magazine*, Spring 2003: 49-56.

Kaufman, L. and Rousseeuw, P.J. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: Wiley.

Kay, J. 1990. Identifying the Strategic Market. *Business Strategy Review*, 1(1): 2-24.

Kelly, C. and Rice, J. 1990. Monotone Smoothing with Application to Dose-Response Curves and the Assessment of Synergism. *Biometrics*, 46: 1071-1085.

Keyte, J.A. and Stoll, N.R. 2004. Markets? We Don't Need No Stinking Markets! The F.T.C. And Market Definition. *Antitrust Bulletin*, Fall 2004: 593-632.

Kovacic, W.E. and Shapiro, C. 2000. Antitrust Policy: A Century of Economic and Legal Thinking. *Journal of Economic Perspectives*, 14(1): 43-60.

Kuhn, T.S. 1970. Logic of Discovery or Psychology of Research?, in M. Curd and J. A. Cover (eds.). *Philosophy of Science: The Central Issues*. London: W.W. Norton & Company.

Kwiatkowski, D., Phillips, P., Schmidt, P. and Shin, Y. 1992. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics*, 54(1-3): 159-178.

Lakatos, I. 1977. Science and Pseudoscience, in M. Curd and J. A. Cover (eds.). *Philosophy of Science: The Central Issues*. London: W.W. Norton & Company.

Lancaster, K. 1966. A New Approach to Consumer Theory. *The Journal of Political Economy*, 74(2): 132-157.

Lancaster, K. 1971. *Consumer Demand*. New York: Columbia University Press.

Lance, G.N. and Williams, W.T. 1967. A General Theory of Classificatory Sorting Strategies: 1 Hierarchical Systems. *The Computer Journal*, 9(4): 373-380.

Leach, D.F. 1992. Absolute Vs Relative Concentration in Manufacturing Industry: 1972-1985. *South African Journal of Economics*, 60(4): 386-399.

Leamer, E.E. 2009. *Macroeconomic Patterns and Stories*. Berlin: Springer-Verlag.

Levin, A., Lin, C. and Chu, C.J. 2002. Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of Econometrics*, 108(1): 1-24.

Lexecon 1994. Beyond Argument: Defining Relevant Markets. *Competition Memo*.

Lexecon 2003. *Introduction to Quantitative Techniques in Competition Analysis*. [Online]. Available: http://www.crai.com/ecp/assets/quantitative_techniques.pdf.

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.

Longino, H.E. 1998. Values and Objectivity, in M. Curd and J. A. Cover (eds.). *Philosophy of Science: The Central Issues*. New York: W.W. Norton & Company: 170-191.

Lopez, J.H. 1997. The Power of the A.D.F. Test. *Economics Letters*, 57(1): 5-10.

Lorr, M. 1983. *Cluster Analysis for Social Scientists*. London: Jossey-Bass.

Maddala, G.S. and Kim, I.-M. 1998. *Unit Roots, Cointegration, and Structural Change*. Cambridge UK: Cambridge University Press.

Mandela Institute 2011. Case Study One on Mergers - Kaya F.M. *Market Definition in Competition Law*. Johannesburg, 17 March 2011.

Manne, G.A. and Wright, J.D. 2010. Innovation and the Limits of Antitrust. *Journal of Competition Law and Economics*, 6(1): 153-202.

Manski, C.F. 1995. *Identification Problems in the Social Sciences*. Cambridge MA: Harvard University Press.

Markovits, R. 2002. On the Inevitable Arbitrariness of Market Definitions. *Antitrust Bulletin*, Winter 2002: 571-601.

Marshall, A. 1920. *Principles of Economics*. London: Macmillan.

Massey, P. 2000. Market Definition and Market Power in Competition Analysis: Some Practical Issues. *The Economic and Social Review,* 31(4): 309-328.

Mawson, N. 2005. Transporter Increases Vehicle Size as Milk Demand Soars. *Engineering News*. Johannesburg, Creamer Media.

McCarthy, T.R. and Thomas, S.J. 2003. Geographic Market Issues in Hospital Mergers, National Economic Research Associates.

McDermott, C.J. and Scott, A. 2000. Concordance in Business Cycles. *Working Paper No. 00/37*, National Bureau of Economic Research.

McFadden, D. 1981. Econometric Models of Probabilistic Choice, in C. F. Manski and D. McFadden (eds.). *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, Mass: MIT Press.

McManus, D.A. 1991. Who Invented Local Power Analysis? *Econometric Theory*, 7(2): 265-268.

Mncube, L., Khumalo, J., Mokolo, R. and Nijisane, Y. 2008. Use of Price Correlation and Stationarity Analysis in Market Definition - Lessons from a Recent Merger. Unpublished paper delivered at the Second Annual Competition Commission, Competition Tribunal and Mandela Institute Conference on Competition Law, Economics and Policy in South Africa. University of the Witwatersrand, Johannesburg.

Mncube, L. and Ratshisusu, H. 2010. On Merger Simulation and Its Potential Role in South African Merger Control. *South African Journal of Economic and Management Sciences*, 13(1): 62-75.

Morgan, B.J.T. and Ray, A.P.G. 1995. Non-Uniqueness and Inversions in Cluster Analysis. *Applied Statistics*, 44(1): 117-134.

Motta, M. 2004. *Competition Policy: Theory and Practice.* New York: Cambridge University Press.

Narayan, P.K. 2005. The Saving and Investment Nexus for China: Evidence from Cointegration Tests. *Applied Economics*, 37(17): 1970-1990.

Neven, D.J. 2006. Competition Economics and Antitrust in Europe. *Economic Policy*, 21(48): 741-791.

Nevo, A. 2001. Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*, 69(2): 307-342.

Neyman, J. 1937. "Smooth" Test for Goodness of Fit. *Skandinavisk Aktuarietiskrift*, 20: 149-199.

Ng, S. and Perron, P. 1995. Unit Root Tests in A.R.M.A. Models with Data Dependent Methods for the Selection of the Truncation Lag. *Journal of the American Statistical Association*, 90(429): 268-281.

Ng, S. and Perron, P. 2001. Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica*, 69(6): 1519-1554.

Niels, G. and Jenkins, H. 2005. Reform of Article 82: Where the Link between Dominance and Effects Breaks Down. *European Competition Law Review*, 26(11): 605-610.

O'Brien, D.P. and Wickelgren, A.L. 2003. A Critical Analysis of Critical Loss Analysis. *Antitrust Law Journal*, 71: 161-184.

Office of Fair Trading 2001. The Role of Market Definition in Monopoly and Dominance Inquiries. *Economic Discussion Paper 2,* London.

Office of Fair Trading and Competition Commission 2010. Merger Assessment Guidelines. London.

Organisation for Economic Co-operation and Development 2003. *Competition Law and Policy in South Africa: An O.E.C.D. Peer Review*. Paris: Organisation for Economic Co-operation and Development.

Pagan, A.R. 1989. 20 Years After: Econometrics 1966-1986, in B. Cornet and H. Tulkens (eds.). *Contributions to Operations Research and Econometrics*. Cambridge, MA: MIT Press.

Pagan, A.R. 2003. An Examination of Some Tools for Macro-Econometric Model Building. Unpublished paper delivered at the Eighth African Econometrics Conference. Stellenbosch.

Perron, P. 1991. Test Consistency with Varying Sampling Frequency. *Econometric Theory*, 7(3): 341-368.

Perron, P. and Ng, S. 1996. Useful Modifications to Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties. *Review of Economic Studies*, 63(3): 435-465.

Perron, P. and Qu, Z. 2007. A Simple Modification to Improve the Finite Sample Properties of Ng and Perron's Unit Root Tests. *Economics Letters*, 94(1): 12-19.

Pesaran, M.H. 2007. A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence. *Journal of Applied Econometrics*, 22(2): 265-312.

Pesaran, M.H., Shin, Y. and Smith., R.J. 2001. Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics*, 16(3): 289-326.

Pesaran, M.H., Smith, R.P., Yamagata, T. and Hvozdyk, L. 2009. Pairwise Tests of Purchasing Power Parity. *Econometric Reviews*, 28(6): 495-521.

Pesaran, M.H. and Timmermann, A. 2005. Real-Time Econometrics. *Econometric Theory*, 21(1): 212-231.

Petrin, A. 2002. Quantifying the Benefits of New Products: The Case of the Minivan. *Journal of Political Economy*, 110(4): 705-729.

Popper, K.R. 1963. Conjectures and Refutations, in M. Curd and J. A. Cover (eds.). *Philosophy of Science: The Central Issues*. London: W.W. Norton & Company.

Porter, M.A., Onnela, J.-P. and Mucha, P.J. 2009. Communities in Networks. *Notices of the American Mathematical Society*, 56(9): 1082-1097.

R Development Core Team 2009. R: A Language and Environment for Statistical Computing. Vienna, Austria, R Foundation for Statistical Computing.

Reekie, W.D. 1999. The Competition Act, 1998: An Economic Perspective. *South African Journal of Economics*, 67(2): 257-288.

Republic of South Africa 1998. *Competition Act, No 89 of 1998*. Pretoria: Government Printer.

Roberts, S. 2004. The Role for Competition Policy in Economic Development: The South African Experience. *Development Southern Africa*, 21(1): 1-17.

Roberts, S., Klaaren, J. and Moodaliyar, K. 2008. Introduction to Special Section on Competition Law and Economics. *South African Journal of Economic and Management Sciences*, 11(3): 247-248.

Roeller, L.-H. and Stehmann, O. 2005. The Year 2005 at D.G. Competition: The Trend Towards a More Effects-Based Approach. *Review of Industrial Organization*, 29(4): 281-304.

Said, S.E. and Dickey, D.A. 1984. Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, 71: 599-607.

Salmon, W.C. 1998. Confirmation and Relevance: Bayesian Approaches, in M. Curd and J. A. Cover (eds.). *Philosophy of Science: The Central Issues*. New York: W.W. Norton & Company.

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.

Sargan, J. and Bhargava, A. 1983. Testing for Residuals from Least Squares Regression for Being Generated by Gaussian Random Walk. *Econometrica*, 51(1): 153-174.

Scheffman, D.T. and Spiller, P.T. 1987. Geographic Market Definition under the U.S. Department of Justice Meger Guidelines. *Journal of Law and Economics*, 30(April 1987): 123-147.

Schumpeter, J.A. 1934. *The Theory of Economic Development*. Cambridge, MA Harvard University Press.

Seo, M.H. 2005. Improving Unit Root Tests by a Generalization of the Autoregressive Spectral Density Estimator at Frequency Zero. *Unpublished manuscript*, Department of Economics, London School of Economics.

Shimodaira, H. 2002. An Approximately Unbiased Test of Phylogenetic Tree Selection. *Systematic Biology*, 51: 492–508.

Shimodaira, H. 2004. Approximately Unbiased Tests of Regions Using Multistep-Multiscale Bootstrap Resampling. *The Annals of Statistics*, 32(6): 2616-2641.

Silk, A.J., Klein, L.R. and Berndt, E.R. 2002. Intermedia Substitutability and Market Demand by National Advertisers. *Review of Industrial Organization*, 20(4): 323-348.

Simon, D. 2004. A Third View of the Black Box: Cognitive Coherence in Legal Decision Making. *The University of Chicago Law Review*, 71(2): 511-586.

Simon, H.A. 1979. Rational Decision-Making in Business Organizations. *The American Economic Review*, 69(4): 493-513.

Sims, C.A. 2001. Comment on 'Evolving Post-War War Ii U.S. Inflation Dynamics. *NBER Macroeconomics Annual*, 16: 373-379.

Slade, M.E. 1986. Exogeneity Tests of Market Boundaries Applied to Petroleum Products. *Journal of Industrial Economics*, 34(3): 291-303.

Smit, M.R. 1999. Competition Policy: South African Literature, Researchers and Data. *TIPS Working Paper No 12*, Trade and Industrial Policy Secretariat.

Smith, A. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Methuen and Company (Limited).

Smith, V.L. 2008. *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge University Press.

Sokal, R.R. and Michener, C.D. 1958. A Statistical Method for Evaluating Systematic Relationships. *University of Kansas Scientific Bulletin*, 38: 1409-1438.

South African Advertising Research Foundation 2005. Radio Audience Measurement Survey.

Spanos, A. 2000. Revisiting Data Mining: 'Hunting' with or without a License. *Journal of Economic Methodology*, 7(2): 231-264.

Stigler, G.J. and Sherwin, R.A. 1985. The Extent of the Market. *Journal of Law and Economics*, 30: 123-147.

Stock, J.H. 2001. Comment on 'Evolving Post-World War Ii U.S. Inflation Dynamics'. *NBER Macroeconomics Annual*, 16: 379-387.

Sutherland, P. and Kemp, K. 2000. *Competition Law of South Africa*. Durban: Butterworths.

Suzuki, R. and Shimodaira, H. 2006. Pvclust: An R Package for Assessing the Uncertainty in Hierarchical Clustering. *Bioinformatics*, 22(12): 1540-1542.

Suzuki, R. and Shimodaira, H. 2009. Pvclust: Hierarchical Clustering with P-Values Via Multiscale Bootstrap Resampling.

The Cambridge Dictionary of Philosophy. 1995. Cambridge, UK: Cambridge University Press.

Theron, N.M. 2001. The Economics of Competition Policy: Merger Analysis in South Africa. *South African Journal of Economics*, 69(4): 614-658.

Theron, N.M. 2010. *Market Definition in Merger Enquiries - Lessons to Be Learnt from the Acquisition of Kaya Fm by Primedia*. *Econex Research Note 17*. [Online]. Available: http://www.econex.co.za/images/stories/ECONEX_researchnote_17.pdf.

Theron, N.M. and Boshoff, W.H. 2006. Vertical Integration in South African Telecommunications: A Competition Analysis. *South African Journal of Economics*, 74(3): 1-18.

Theron, N.M. and Boshoff, W.H. 2011. When Do Vertical Restraints Harm Competition? The Economics-Based Approach and Its Application in the Batsa Case. *South African Journal of Economics*, forthcoming.

Theron, N.M. and Van Eeden, J. 2011. *Asymmetric Mobile Termination Rates in South Africa*. *Econex Research Note 21*. [Online]. Available: http://www.econex.co.za/images/stories/ECONEX_researchnote_21.pdf.

Theron, N.M. and Van Eeden, J. 2011. *Two-Sided Markets: An Application to Mobile Termination Markets in South Africa*. *Econex Research Article 15*. [Online]. Available: http://www.econex.co.za.

Tibshirani, R., Walther, G. and Hastie, T. 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.

Trade Practices Commission 1994. Market Definition and Competition Issues in Commercial Broadcast Radio. Canberra, Australian Government Publishing Service.

Tukey, J.W. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley Publishing Company.

Turner, P. 2006. Response Surfaces for an *F*-Test for Cointegration. *Applied Economics Letters*, 2006(13).

United Kingdom Competition Commission 2008. *The Supply of Groceries in the U.K. Market Investigation*. London.

United States Department of Justice and Federal Trade Commission 1992. Horizontal Merger Guidelines. Washington DC.

United States Department of Justice and Federal Trade Commission 2010. Horizontal Merger Guidelines. Washington DC.

Urbakh, V.U. 1971. Linear Discriminant Analysis: Loss of Discriminating Power When a Variate Is Omitted. *Biometrics*, 27(3): 531-534.

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.

Vogelsang, T.J. 1999. Two Simple Procedures for Testing for a Unit Root When There Are Additive Outliers. *Journal of Time Series Analysis*, 20: 237-252.

Waldfogel, J. 2002. Consumer Substitution among Media. *Media Bureau Staff Research Papers 2002-2003*.

Walesiak, M. and Dudek, A. 2009. Clustersim: Searching for Optimal Clustering Procedure for a Data Set.

Werden, G.J. and Froeb, L.M. 1993. Correlation, Causality and All That Jazz: The Inherent Shortcomings of Price Tests for Antitrust Market Delineation. *Review of Industrial Organization*, 8: 329-353.

White, H. 2000. A Reality Check for Data Snooping. *Econometrica*, 68(5): 1097-1126.

Williamson, O.E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.

Yule, G.U. 1926. Why Do We Sometimes Get Nonsense Correlations between Time Series? A Study in Sampling and the Nature of Time Series. *Journal of the Royal Statistical Society*, 89(1): 1-64.

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 RAMS 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, 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

Appendix B

1-standard Deviation Rule for Gap Statistic

Figure B1: Standard deviation rule for gap statistic based on average linkage method

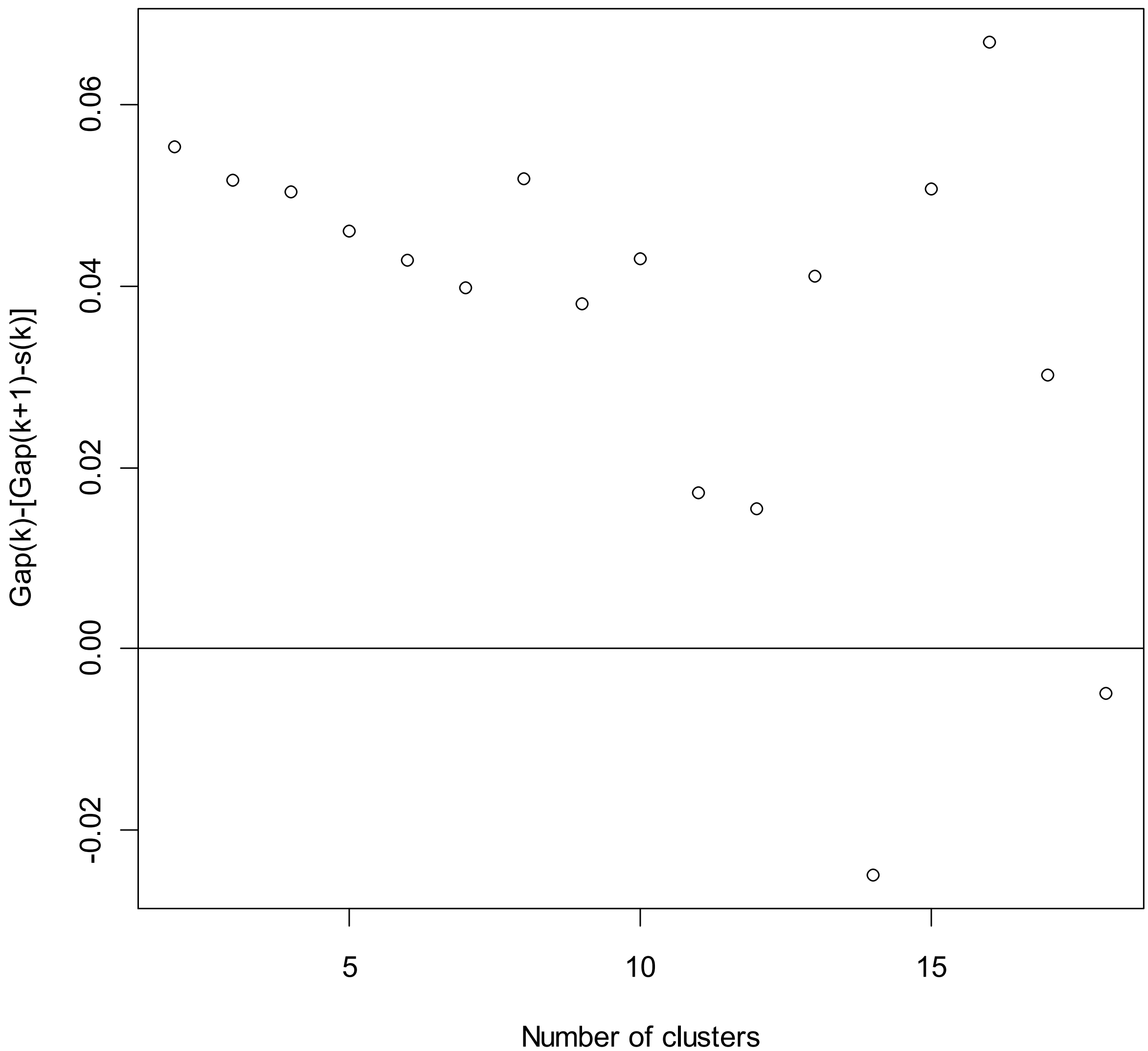


Figure B2: Standard deviation rule for gap statistic based on single linkage method

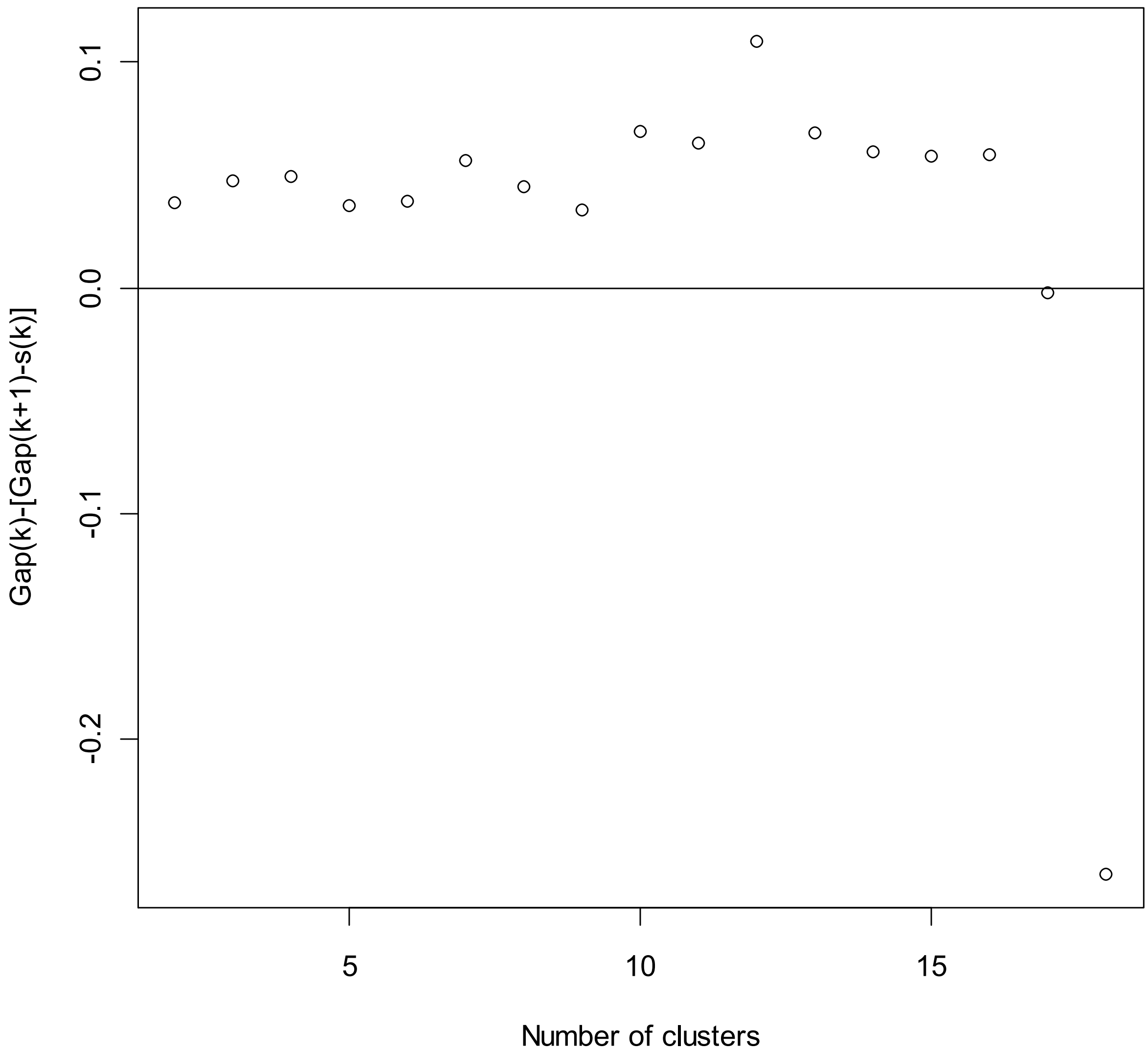
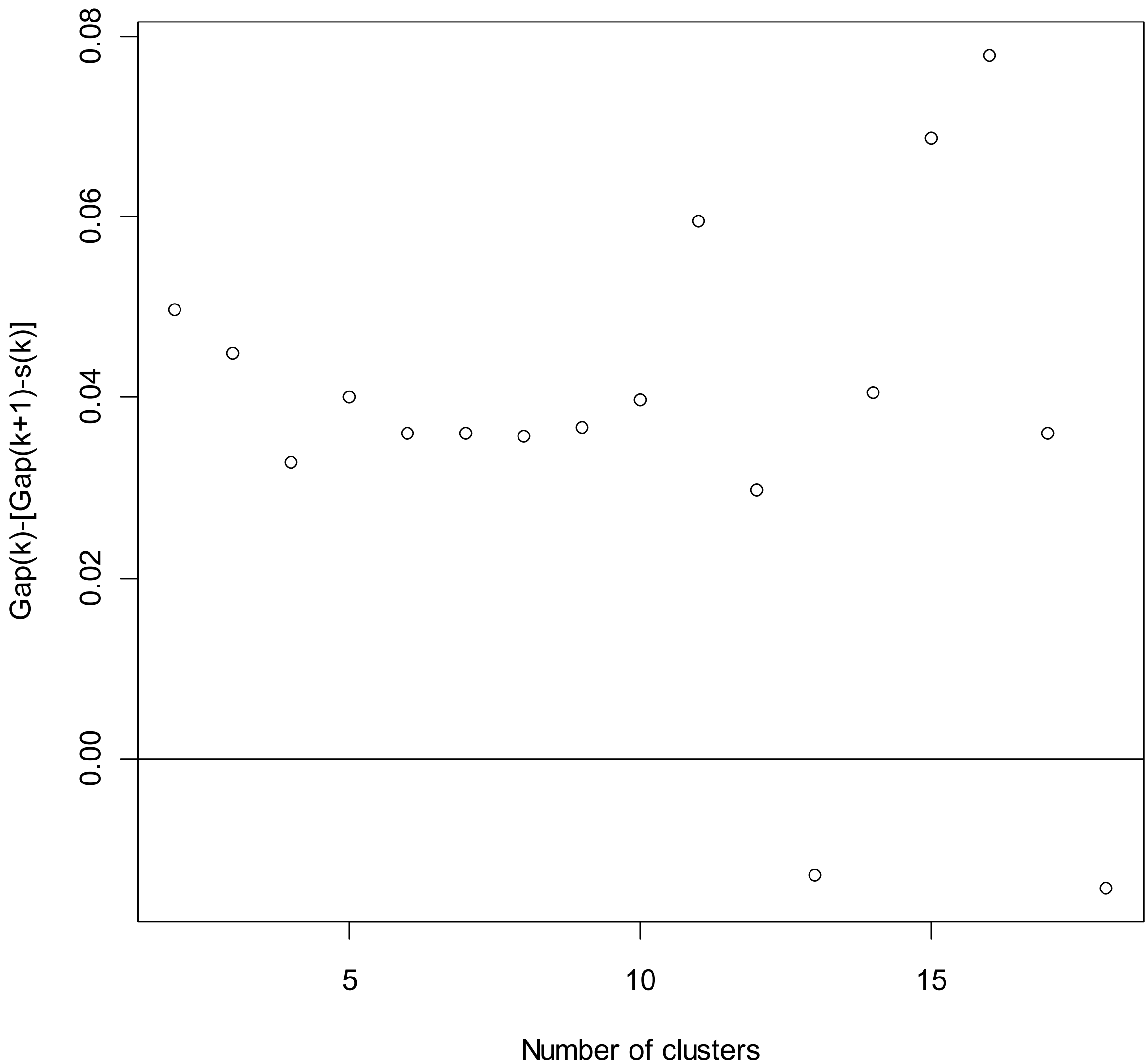


Figure B3: Standard deviation rule for gap statistic based on complete linkage method



Appendix C
Lag Length Selection Results
(Sample Period January 2002 – December 2005)

Table C1: Results for the Western and Eastern Cape system

Lag	Log-likelihood	LR	FPE	AIC	SC	HQ
0	139.0862	NA	5.01e-06	-6.527913	-6.445167	-6.497583
1	201.4908	115.8943	3.11e-07*	-9.309085*	-9.060846*	-9.218095*
2	203.6061	3.726943	3.41e-07	-9.219337	-8.805606	-9.067688
3	204.7154	1.848929	3.92e-07	-9.081687	-8.502464	-8.869379
4	207.4966	4.370392	4.19e-07	-9.023647	-8.278931	-8.750679
5	208.0705	0.847148	4.99e-07	-8.860498	-7.950290	-8.526871
6	215.1342	9.754745*	4.38e-07	-9.006392	-7.930692	-8.612106

* indicates lag order selected by the criterion

LR: sequential modified likelihood ratio test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

Table C2: Results for the Southern and Eastern Cape system

Lag	Log-likelihood	LR	FPE	AIC	SC	HQ
0	177.0435	NA	8.22e-07	-8.335406	-8.252660	-8.305076
1	227.5362	93.77217	8.99e-08	-10.54934	-10.30111*	-10.45836*
2	227.9817	0.784797	1.07e-07	-10.38008	-9.966348	-10.22843
3	236.7005	14.53142*	8.55e-08*	-10.60479*	-10.02556	-10.39248
4	236.9677	0.419878	1.03e-07	-10.42703	-9.682318	-10.15407
5	238.6363	2.463129	1.16e-07	-10.31601	-9.405805	-9.982386
6	242.6799	5.584005	1.18e-07	-10.31809	-9.242389	-9.923802

Table C3: Results for the Western and Southern Cape system

Lag	Log-likelihood	LR	FPE	AIC	SC	HQ
0	158.1860	NA	2.02e-06	-7.437427	-7.354681	-7.407097
1	213.2576	102.2758*	1.77e-07*	-9.869407*	-9.621169*	-9.778418*
2	213.8812	1.098835	2.09e-07	-9.708630	-9.294899	-9.556981
3	214.7589	1.462788	2.43e-07	-9.559947	-8.980724	-9.347639
4	215.9921	1.937963	2.79e-07	-9.428197	-8.683482	-9.155230
5	218.1462	3.179815	3.09e-07	-9.340296	-8.430088	-9.006669
6	220.5580	3.330510	3.39e-07	-9.264665	-8.188964	-8.870378

Appendix D

Price Ratio Unit Root Tests

(Sample Period January 2002 – December 2005)

Table D1: Augmented Dickey-Fuller (1979) tests for January 2002 – December 2005

Lag	*t*-stat	p-value
Western : Eastern		
1	-2.46	0.35
2	-2.28	0.44
3	-1.77	0.70
4	-1.72	0.72
5	-0.92	0.94
6	-0.75	0.96
Southern : Eastern		
1	-2.98	0.15
2	-1.89	0.64
3	-2.01	0.58
4	-1.85	0.66
5	-1.35	0.86
6	-1.70	0.73
Western : Southern		
1	-1.63	0.77
2	-1.42	0.84
3	-1.36	0.86
4	-1.40	0.84
5	-1.50	0.81
6	-1.38	0.85

Note: *** Reject at 1% ** Reject at 5% * Reject at 10%

Table D2: Ng-Perron (1996, 2001) tests for January 2002 – December 2005

Lag	MZ_a	MZ_t	MSB	MP_T
Western : Eastern				
1	-11.38	-2.37	0.21	8.10
2	-11.77	-2.41	0.20	7.83
3	-6.34	-1.76	0.28	14.36
4	-6.47	-1.78	0.27	14.08
5	-3.56	-1.30	0.37	25.13
6	-3.71	-1.33	0.36	24.14
Southern : Eastern				
1	-16.07	-2.79	0.17	5.94
2	-7.74	-1.90	0.25	11.92
3	-11.96	-2.39	0.20	7.90
4	-12.72	-2.47	0.19	7.44
5	-7.76	-1.91	0.25	11.90
6	-28.75***	-3.76***	0.13***	3.37***
Western : Southern				
1	-6.31	-1.72	0.27	14.43
2	-5.44	-1.59	0.29	16.57
3	-5.02	-1.52	0.30	17.86
4	-7.71	-1.91	0.25	11.93
5	-15.98*	-2.79*	0.17*	5.91*
6	-20.84**	-3.20**	0.15**	4.56**

Note: *** Reject at 1% ** Reject at 5% * Reject at 10%