Inferring Market Definitions and Competition Groups From Empirically-Estimated Demand Systems: A Practitioner's Guide*

Assaf Kovo † Alon Eizenberg‡

May 2017

Abstract

This paper addresses several gaps in the economics and antitrust literature concerning the task of formally defining the boundaries of a market. First, we provide a formal treatment of the "submarket" concept referred to in seminal court rulings. To that end, we define a concept termed "competition group" (CG) and formally relate it to the familiar SSNIP test. Second, we provide a methodology for performing the SSNIP test and constructing CGs using an empirically estimated demand system, addressing several practical difficulties that have not been discussed in the extant literature. Our approach highlights the practical virtues of the nested logit model of demand in this context. Finally, we provide a detailed description of the application of this methodology by the Israeli Antitrust Authority in its investigation of the country's dairy market.

Keywords: Market definition, Differentiated product markets; Nested logit models, Competition policy, Antitrust. **JEL Classification:** D12, L40, K21.

*This paper presents our personal views, not necessarily shared by the Israeli Antitrust Authority.

†Department of Economics, Hebrew University of Jerusalem and Israeli Antitrust Authority, assaf.kovo@mail.huji.ac.il

‡Department of Economics, Hebrew University of Jerusalem and CEPR, alon.eizenberg@mail.huji.ac.il

1 Introduction

Defining the market to which a given product belongs is a fundamental task in the study and regulation of markets. Practitioners define product markets to accomplish a variety of goals, a natural one being merger analysis: determining whether product B belongs in product A's market is a key step in evaluating the likely consequences of a merger between the firms that produce these two products.[1] Classifying a firm's conduct as "monopolization" is another task that requires determining the market in which the firm competes, and whether the said firm is a "price maker" rather than a "price taker" to the point it has monopoly power.[2] [3]

A key ruling in U.S. law regarding criteria for market definition is Brown Shoe Co. v. United states (hereafter "*Brown Shoe*") stating that:[4]

> "The outer boundaries of a product market are determined by the reasonable interchangeability of use or the cross-elasticity of demand between the product itself and substitutes for it. However, within this broad market, well-defined submarkets may exist which, in themselves, constitute product markets for antitrust purposes. The boundaries of such a submarket may be determined by examining such practical indicia as industry or public recognition of the submarket as a separate economic entity, the product's peculiar characteristics and uses, unique production facilities, distinct customers, distinct prices, sensitivity to price changes, and specialized vendors."
>
> 370 U.S. at 325 (footnotes and citation omitted).

Two key issues are embedded in this quote. The first is the introduction of the term "submarkets", while the second involves the importance of the cross-elasticity of demand between products in defining the boundaries of the market. Our paper wishes to shed light on both issues by discussing a methodological approach that allows one to empirically estimate a model of consumer demand, and then use the estimated model in order to evaluate the extent to which particular partitions of the market into well-defined "competition groups" are consistent with the data, where competition groups represent our alternative to the concept of submarkets.

Our paper has three components: first, we model the concept of competition groups, formally relate it to the familiar SSNIP test and compare it to the concept of submarkets. We then propose the estimation of a particular model—the nested logit model of differentiated product demand—as a practical tool for defining competition groups. Finally, we demonstrate via a

[1] See for example United States v. Mrs. Smith's Pie Co. 440 F. Supp. 220 (1976)

[2] See for example United States v. Alcoa 148 F.2d 416 (2d Cir. 1945), United States v. Microsoft Corporation 253 F.3d 34 (D.C. Cir. 2001), or Eastman Kodak Company v. Image Technical Services, Inc. 504 U.S. 451 (1992).

[3] See U.S. Department of Justice, Competition and Monopoly: Single-Firm Conduct Under Section 2 of the Sherman Act (2008).

[4] 370 U.S. 294 (1962).

specific example: the Israeli Antitrust Authority's (IAA) investigation of the country's dairy-based spreads market.[5]

Submarkets. The notion of submarkets is far from being self-explanatory, as evident in the literature. On the one hand, critics like Baker (2000) claim that "there are no submarkets, only markets"; while on the other hand, *Brown Shoe* is still "the most cited Supreme Court case when it comes to defining product markets" (Pettit 2008). Some of the difficulty with this concept may be attributed to its apparent incompatibility with the SSNIP test, the common market definition tool.[6] The SSNIP test is applied by focusing on a specific product (hereafter the "focal" product), and determining the set of products that belong in the focal product's market. This definition is inherently asymmetric: it implies that product A may belong in product B's market, even though the converse is not true. While this asymmetry does not create any theoretical difficulty, it appears to contradict the spirit of submarket definition, which is inherently symmetric: A and B are either members of the same submarket, or not.

Our suggested definition of competition groups bridges this gap by maintaining the symmetry property while maintaining a direct link to the SSNIP test. The definition of competition groups can be attractive in many important, practical applications. Mainly, it creates a differentiation between a particular *subgroup of products* that demonstrate high degree of substitutability between themselves and therefore are candidates to be included together in a specific well-defined market, and products outside of that subgroup that demonstrate low degree of substitutability to the products of that subgroup and so are not likely to be part of any well-defined market containing products from it. This approach may sometimes be more helpful than defining a market over each of (potentially) hundreds of products (treating each as the "focal" product, one at a time).

In the automobile market, for instance, the SSNIP test may allow us to determine that cars B, C, and D belong in car A's market in a given point in time. Following the replacement of car D by its more advanced version D', a frequent event in the industry, it will not be immediately clear that the previous finding would remain valid. In practice, it will not be reasonable for the antitrust authority to collect data and re-perform the market definition in response to such frequent modifications in the set of products offered to consumers. It may, therefore, be of interest to develop a procedure that determines whether a well-defined subgroup of products constitutes a distinct market, since such a finding would constitute a more stable statement concerning market segmentation. This subgroup might be, for example, "all compact cars", or "all compact hybrid cars" depending on the qualitative findings. The practical indicia described

[5]Dairy-based spreads are in essence, various cheeses. The term "spreads" is used as part of the conceptual framework of the investigation, discussed in detail in Section 4.

[6]Small but Significant and Non-transitory Increase in Price. According to the American Horizontal Merger Guidelines (2010), (hereafter "*AHMG*"), the SSNIP is employed as a methodological tool for performing the hypothetical monopolist test which will be characterized in Sections 2 and 3.

in *Brown Shoe*, already a part of qualitative analyses carried out by competition authorities, should therefore play a major role in defining product subgroups.

Using empirical estimates of demand elasticities in the definition of markets and competition groups. The court decision in Federal Trade Commission v. Whole Foods Market, Inc. emphasized the importance of cross-elasticity of demand in defining the boundaries of the market, and specifically the role of the SSNIP test, when it favored economic evidence regarding consumers' reactions to a price increase over non-economic practical indicia similar to those described in *Brown Shoe*.[7]

Common versions of the SSNIP test, as found in the literature, do not utilize estimated demand elasticities. Rather, the procedures use other quantities that are often available to competition authorities: measures of products' sales and margins. These procedures use these quantities to make statements about profitability that are typically only valid under very stringent assumptions, such as symmetric Cournot competition. An estimated demand model can offer an alternative set of assumptions under which the SSNIP test can be performed.

The empirical literature in Industrial Organization highlights the role of empirical estimation of consumer demand. This literature offers models of differentiated product demand systems, as well as estimation techniques that enable one to make inference on key primitives: the parameters that govern the distribution of consumer preferences (see Bresnahan (1981, 1987), Berry (1994), Berry et al. (1995, BLP), Nevo (2001)). Such models inform us about the extent of correlation of a consumer's valuation across products that share certain observed characteristics, or belong in particular subgroups of products. As a consequence, these models provide a picture of market segmentation: to what extent are products "close" or "distant" substitutes. Indeed, this literature cites market definition as one of the main motivations to estimate differentiated product demand systems (e.g., Nevo 2000).

While the literature on demand estimation is well-established, *the practical definition of product markets using demand estimates* receives much less attention in the extant empirical literature. Specifically, certain gaps exist in this literature regarding the practical application of the SSNIP test. Our paper attempts to help close some of these gaps, while making the extra step of moving from application of SSNIP to the definition of competition groups. Our approach highlights the role of a particular econometric model—the nested logit model of differentiated product demand—as a practical tool for this task.

The nested logit model allows the researcher to a-priori classify products into groups based on institutional details (e.g., the qualitative analysis often performed by antitrust authorities based on discussions with industry insiders). The model can then be estimated using widely-available product-level data. These estimates, along with our suggested application of the SSNIP test, allow the researcher to assess the extent to which the a-priori market segmentation is consistent

[7] 548 F.3d 1028 (D.C. Cir. 2008).

with the data. We further discuss below how application of the SSNIP test can be used to formally define competition groups.

Scope and caveats. We view the role of the quantitative analysis described in this paper as one which is complementary to the more typical, qualitative analysis often applied by practitioners. This stems from a number of reasons. First, the procedure begins with an a-priori market segmentation, and then investigates the extent to which the data are consistent with this segmentation, while comparing the performance of the suggested segmentation relative to some natural alternatives. Our suggestion is to use the practitioner's qualitative assessment of the market segmentation as the a-priori segmentation applied in the quantitative analysis. To the extent that this segmentation is not rejected by the data, and appears superior to natural alternatives, this procedure can provide support to the qualitative assessment.

Another possible use of the methodology could be to help practitioners select among several plausible hypotheses regarding market segmentation, as an intermediate analytical step. The flexibility and relative simplicity of the estimation procedure, which is carried out by linear regressions, allows the researcher to experiment with several models of segmentations and to compare their performance in a manner that may help augment her intuition regarding the empirical magnitudes of the underlying substitution patterns. This may be of particular value if different, plausible hypotheses are presented to the practitioner (e.g., conversations with some market participants may suggest that brand is the strongest segmenting factor, while conversations with others may suggest that some other product characteristic has a dominant effect in segmenting the market). The methods described in this paper, therefore, are not meant to provide a statistical black box that selects the correct market segmentation out of a universe of all possible segmentation. Rather, it is a quantitative tool that can help practitioners bring plausible hypotheses regarding segmentation to the data in search of some supporting, or contradicting, empirical evidence.

Finally, it should be noted that the application of such models depends heavily on the availability of appropriate data and the resources to process them. The data should include some variation with respect to prices while consumer preferences remain stable. Also, in markets where firms compete mostly on quality, such models may prove to be difficult to implement.

The paper proceeds as follows: after a brief literature review offered below, Section 2 provides a formal model: first, we formally define the concept of competition groups, viewed as a partition of a (potentially large) set of goods. Second, we spell out the assumptions of the nested-logit model of consumer demand. Section 3 discusses the practical implementation: estimation of the nested logit model, application of SSNIP tests for individual (focal) products, and then finally using the outcome of these tests to validate the a-priori partition of goods into competition groups. Section 4 demonstrates via the application to the Israeli dairy-based spreads market, while Section 5 offers concluding remarks.

Literature. The SSNIP is used when performing a critical loss analysis ("CLA"), a term coined by Harris and Simons (1989). CLA is an empirical method that investigates the intensity of competitive interaction (Hüschelrath 2009). It compares the predicted percentage of units a firm will lose as a result of a SSNIP ("actual loss"), to the maximal percentage of lost units the firm can sustain for the SSNIP to be profitable ("critical loss"), in order to determine the borders of the market. The result of the test relies heavily on primitives such as products' own and cross elasticities, products' markups, the type of competition between firms and each firm's product portfolio. The IO literature, and specifically the antitrust literature, is concerned with constructing formulae that apply to different compositions of these primitives. Katz and Shapiro (2002, KS) and Daljord et al. (2008) discuss how to properly apply cross-elasticities into the critical loss; Farrell and Shapiro (2007) discuss how to account for rivals' price responses; Moresi et al. (2008) discuss the implementation of the SSNIP test with multi-product firms; and O'Brien and Wickelgren (2003) show that firms with larger pre-merger margins will raise prices more than firms with smaller margins.

Most papers assume symmetric products and profit maximization for ease of computation. Scheffman and Simons (2003) argued against the latter and the common use of the Lerner equation that originated from it while Daljord (2009) addressed the former with an application of the asymmetric case to the computation of the critical loss. Also, Daljord and Sørgard (2011) showed that product asymmetry can lead to broader market definitions under a uniform SSNIP test than under a single-product SSNIP test.[8]

Finally, it is important to note that models of differentiated product demand, which allow us to overcome most of the aforementioned issues, can be used in various markets. Examples are numerous and range from automobiles (BLP 1995, Verboven 1996), telecommunications (Werden and Froeb 1994), computers (Eizenberg 2014), R&D decisions (Kaiser 2002), analgesics (Björnerstedt and Verboven 2013) and cereals (Nevo 2001). Some of these papers (notably BLP, Nevo 2001 and Villas-Boas 2007) show deduced markups given estimated demand systems. While antitrust authorities have the prerogative to obtain data directly from firms that are the subject of an investigation, including data regarding costs and markups, such exercises may still provide valuable information, as we discuss below.

2 A Formal Model

This section proceeds in two subsections. Subsection 2.1 provides basic definitions of products and markets, as well as of the SSNIP test. This subsection also describes the variables that we assume throughout to be observed by the researcher. Subsection 2.2 then follows with a model

[8] The terms "uniform" and "single-product" refer to the number of products that experience a price increase as part of the SSNIP test. We address this issue in Section 3.3.

of differentiated-product demand in which the distribution of consumer preferences serves as the primitive of interest.

2.1 Product and Market Definitions and Observed Data

Product definitions. The definitions here follow Berry (1994) exactly. Consider a set of differentiated products $\mathcal{J}$ indexed by $j = 1, ..., J$ with $J = |\mathcal{J}|$.[9] The set $\mathcal{J}$ can be thought of as a broad group of products that may, potentially and a-priori, include subsets of products that compete with one another. In the context of the automobile market, this set may include all vehicle models that are sold to consumers in the United States within a given year. The goal of our analysis is to determine whether specific subsets of $\mathcal{J}$ constitute well-defined "markets", where the concept of such markets would be precisely defined below.

Two other, closely-related aspects of our analysis are the presence of an "outside option" denoted $j = 0$, and the size of the market denoted M. The importance of an outside option has been discussed by many authors (e.g., Berry 1994). In the absence of such an option, the model would fail to capture substitution toward products about which the researcher has much less precise information. In the automobile example, the "inside goods" $j = 1, ..., J$ may include all new models available for consumer purchase, while the outside option $j = 0$ may capture alternatives such as not owning a car, buying a used car, etc. Failure to account for the outside option may yield unreasonable predictions: for example, an increase in the prices of all inside goods would not imply a drop in the total, combined quantity of all these inside goods.

The size of the market M is defined as the largest potential quantity of units sold of all products $j = 0, 1, ..., J$ (i.e., including the outside option). Our maintained assumption would be that M is observed or assumed, and we shall come back to this important issue in the sections below. We next define *market shares* as follows: for the inside goods $j = 1, ..., J$, product j's market share is given by $s_j = q_j/M$, where q_j is the total quantity sold of this product. The share of the outside option is defined by $s_0 = 1 - \sum_{j=1}^{J} s_j$.

Observed data. We assume throughout that we observe data from a set of markets $t = 1, ...T$. Such markets may be defined in several manners. For instance, they are defined geographically (e.g., a cross-section of automobile markets in European countries as in Verboven (1996), or they may capture a time-series dimension (as in BLP, who study the US automobile market over a period of 20 years, such that $t = 1, ..., 20$).

In each market t, we observe the set of inside products $\mathcal{J}_t$, with products indexed by $j = 0, 1, ..., J_t$ where $J_t = |\mathcal{J}_t|$. We also observe the market size M_t, and the quantity q_{jt} sold of good. We can then easily calculate $s_{jt} = q_{jt}/M_t$ and the outside share $s_{0t} = 1 - \sum_{j=1}^{J_t} s_{jt}$. For each inside good we also observe its price, p_{jt}, and a vector of product characteristics $x_{jt} \in \mathcal{R}^k$.

[9] The notation and several aspects of the conceptual framework introduced in this Section draw on Berry (1994).

In the automobile example, this k-vector of characteristics may include Horse Power, size, fuel efficiency, etc.

Terminology and definitions. The term "market" can have various context dependant meanings. To avoid confusion we use distinct terms for each concept. First, the term "market" itself has already been introduced above and will be used following the framework of Berry (1994). Second, the concept of a "market" as discussed in the AHMG and refers to the antitrust definition of markets will be referred to as "Relevant Market" and will always be associated with a specific product (RM_j). Third, we consider partitions of the products in a given market $j = 1, ..., J_t$ into subsets, and, given a partition that satisfies certain criteria described below, refer to the subsets as "Competition Groups" (CG).

Product $j \in \mathcal{J}_t$'s relevant market RM_j is defined by the narrowest group of products that includes product j and enables the exercise of market power. This relevant market is defined by practitioners using the hypothetical monopolist test (hereafter "*HMT*"). The core concept of this test is that a single-product firm will find it hard to unilaterally and profitably increase its price as consumers will turn to purchase substitute products. As the firm's product portfolio grows larger, this effect is diminished as some of the substitution will be towards other products controlled by the same firm, implying that a price increase can become more profitable.[10]

In the HMT, the question is whether the optimal decision of a profit-maximizing firm constitutes a SSNIP. A price increase of 5% is widely accepted as the threshold SSNIP, therefore the test should identify the narrowest group of products that includes product j for which the optimal pricing scheme, given ownership of a single firm over that group of products, results in a price increase of at least one product in that group by at least 5%. However, for reasons discussed in Section 3 this definition can prove to be difficult to implement. Thus, we adopt a slightly different definition based on KS. Next, we formulate the two definitions and discuss the differences between them.

First, we offer some general notations. We denote by p^0_{jt} the base (observed) price of each of the inside goods in a specific market $j = 1, ..., J_t$ so that $p^0_t = (p^0_{1t}, ..., p^0_{J_t t})$ is the vector of the observed prices in that market.[11] Given three products $i, j, k \in \mathcal{J}_t$, we denote $k \succ_j i$ if product k is a closer substitute to product j than is product i. Moreover, we have that $j \succ_j k$ for each k such that $j \neq k$.[12] Without loss of generality, let us order the indices of all products in a decreasing order of substitutability to some product j and denote by S^j_i the set of the i closest

[10] According to the AHMG: "the test requires that a hypothetical profit-maximizing firm, not subject to price regulation, that was the only present and future seller of those products ("hypothetical monopolist") likely would impose at least a small but significant and non-transitory increase in price ("SSNIP") on at least one product in the market, ... For the purpose of analyzing this issue, the terms of sale of products outside the candidate market are held constant."

[11] We will address the shortcomings of the SSNIP test with respect to the existing type of competition between firms, or lack of it, in Section 3.3

[12] We will address the issue of different ways to formulate substitutability ranking in Section 3.3.

substitutes to product j.[13] For the purpose of defining the SSNIP test we assume that a single firm owns every product in S_i^j while the ownership of other products remains unchanged. We denote by $\Pi_{kt}(p_t)$ the variable profit of each inside good $k \in J_t$, and use $\Pi_{kt}(p_t)^0$ to denote the base variable profit of this product given the observed prices.

Second, we formulate the two SSNIP tests, according to both the AHMG and KS. Following the AHMG, a set of products S_i^j satisfies the SSNIP test if there exists some product $k \in S_i^j$ such that $p_{kt}^* \geq 1.05 \cdot p_{kt}^0$ where $p_{kt}^* \in p_t^*(S_i^j)$ and $p_t^*(S_i^j)$ is the price vector that maximizes $\sum_{k \in S_i^j} \Pi_{kt}(p_t)$, given that $p_{lt} = p_{lt}^0$ for each $l \notin S_i^j$.[14]. In contrast, based on KS, S_i^j satisfies the SSNIP test if for a price vector $\hat{p_t}$ that satisfies $\hat{p_{kt}} = p_{kt}^0$ for every $k \neq j$ (whether $k \in S_i^j$ or $k \notin S_i^j$) and $\hat{p_{jt}} = 1.1 \cdot p_{jt}^0$ it holds that $\sum_{k \in S_i^j} \Pi_{kt}(\hat{p_t}) \geq \sum_{k \in S_i^j} \Pi_{kt}(p_t)^0$. This version of the SSNIP test measures whether a 10% increase in the focal product itself does not lead to a decrease in the profits of the hypothetical monopolist.[15] The obvious difference between the two versions is that the researcher has to calculate the profit maximizing price vector in former version, while in the latter version the examined price change is predetermined. This difference not only makes the latter version significantly easier to compute, but also emphasizes the importance this version assigns to the focal product and its role in determining the borders of its own relevant market. We believe that the version of the SSNIP test based on KS has also the advantage of being more intuitive and easily explained.

Throughout this work we use the terms HMT and SSNIP test interchangeably while referring to the aforementioned procedure.

Definition 1. ***(Relevant Market).*** *Given a set of products $\mathcal{J}_t$ and an observed price vector p_t^0, we define S_i^j to be RM_j if i is the lowest indexed product such that S_i^j satisfies the SSNIP test based on KS.*

The relevant market to some product j is therefore the smallest set of the closest substitutes to that product that satisfy the SSNIP test. We next turn to a key motivation for this paper: effectively determining whether the substitution patterns within a specific subgroup of products are significantly different from those with outside products, so the products in that subgroup can be considered together in any well-defined market. Such subgroups are referred to as Competition Groups (CG). We present our definition for this concept, building on the definition for relevant markets.

Definition 2. ***(Competition Group).*** *A set of products $\mathcal{S} \subseteq \mathcal{J}_t$ is a CG if for every product $j \in \mathcal{S}$ it holds that $RM_j \subseteq \mathcal{S}$.*

[13] Formally we have that each product $a \in S_i^j$ satisfies $a \succ_j i$ and there does not exist a product z such that $z \in \mathcal{J}_t \backslash \{S_i^j\}$ and $z \succ_j i$.

[14] Clearly $p_t^*(S_i^j)$ may change depending on the products portfolio of the firm controlling S_i^j.

[15] Formally, in their work KS discuss a price increase "of at least one of the products." We elaborate on the subject of choosing the product, or products, whose price should be increased in Section 3.3. The definition of the test we present here is in accordance with the methodology described in the empirical example we discuss in Section 4.

In words, a set of products is a CG if the relevant market of each product in that set does not include products out of that set. This definition allows for overlap across CGs (i.e., the same product can belong in multiple CGs), and in the extreme case it might even be that $\mathcal{S} \subset \mathcal{S}'$ where both $\mathcal{S}$ and $\mathcal{S}'$ are CGs. One of the major controversies around the concept of submarkets is that if a set of products constitutes a product market for antitrust purposes, as suggested by Brown Shoe, then it is, by definition, a market and not a submarket. Our definition of CGs is an attempt to tackle this inconsistency by offering a flexible concept that relies on each focal product's market definition.

While the above definition allows for overlapping CGs, it is of practical value to consider cases where the researcher begins with a *partition* of the set of products $j \in \mathcal{J}_t$ into mutually-exclusive subsets, and to consider whether these subsets constitute CGs. This approach loses some of the generality implied by the definition of CGs above, and, in particular, it may be the case that the data will reject a proposed partition due to the misplacement of only several products.[16] Alternatively, it could also be the case that multiple partitions will not be rejected by the data. Just the same, we believe that this is a natural strategy that fits well with the conceptual and empirical issues that arise in this context.

The a-priori partition of $\mathcal{J}_t$ will not be random but rather based on a qualitative analysis performed by the practitioner. In principle, one could develop an automated algorithm that considers many (or all) possible partitions and selects among them the partition(s) that best fit the data according to some criterion (loss function). In our view, this is a less attractive avenue, and we therefore advocate for the construction of candidate CGs based on qualitative economic reasoning, followed by their testing using the data. This approach motivates using the framework developed above as a tool for hypothesis testing and corroboration.[17] This approach also offers a coherent analysis in the event that the partition into CGs is not uniquely defined.

Given a candidate CG $\mathcal{S}$, it might be the case that for a certain product $k \in \mathcal{S}$, RM_k includes products outside of $\mathcal{S}$. According to the definition above, $\mathcal{S}$ does not qualify as a proper CG. Just the same, if s_{kt} is relatively small, the researcher may still be willing to consider $\mathcal{S}$ as a CG. We therefore propose the following refinement to the definition of a CG:

Definition 3. ***(ψ-Competition Group).*** *A set of products $\mathcal{S} \subseteq \mathcal{J}_t$ is a ψ-CG if $\sum_{j \in Y} s_{jt} \geq \psi$, where $j \in Y$ if $j \in \mathcal{S}$ and $RM_j \subseteq \mathcal{S}$.*

2.2 Consumer Demand

This section introduces our demand model. Conceptually, the empirical estimation of this model's parameters would inform us on the nature and magnitude of substitution patterns that char-

[16] We discuss the various ways the data can reject a specific partition in Section 3.2.
[17] See Coate and Fischer (2008) for a discussion of hypothesis testing and corroboration at the FTC.

acterize the underlying demand system. Having identified these patterns, one can proceed by defining an individual product's *Relevant Market* (Definition 1) using the SSNIP test: intuitively, the estimated demand system provides us with an ordering of products according to the extent to which they serve as close substitutes to the focal product. It is this ordering that allows us to perform the SSNIP test. Having defined *Relevant Markets* for each product, we then rely on Definition 3 to define $\psi-$*Competition Groups* that describe the market's segmentation (see Section 3 below for a detailed description of the practical implementation of these steps).

We model consumer demand following the Nested Logit model (McFadden et al. 1978). Berry (1994) shows how to use Cardell's (1997) representation of this model to enable linear estimation methods, incorporate product-level demand errors, and account for price endogeneity using instrumental variables. We adopt that approach (as well as much of the associated notation) here, dropping market indices $t = 1, ..., T$.

The model classifies products $j = 1, ..., J$ into G mutually-exclusive sets, or "nests." The outside option good $j = 0$ is the only member of its own set, leading to a total of $G + 1$ sets. Consumers are assumed to purchase at most one of the products available for sale, or choose the outside option. These choices maximize an individual utility function that captures the consumer's preferences. Consumer i's utility from product $j \in g$ is given by:

$$u_{ij} = x_j\beta + \alpha p_j + \xi_j + \zeta_{ig}(\sigma) + (1 - \sigma)\epsilon_{ij} \tag{1}$$

where the following notation is used: x_j is a vector of product characteristics that are observed by the econometrician (see the discussion in Section 2.1 above). The parameter vector β, to be estimated, captures the utility weights ascribed by consumers to these product characteristics. The price of the product is p_j, implying that the parameter α captures price sensitivity, and is expected to be negative. The random term ξ_j captures utility shifters that are unobserved by the econometrician. Since the observed vector x_j will likely fail to capture all the aspects that affect utility and demand, the inclusion of this term is important. In particular, ξ_j may capture aspects of the good that are difficult to quantify, such as the seller's reputation. While the econometrician does not observe this shifter, firms and consumers are assumed to observe it.

The random term $\zeta_{ig}(\sigma))+(1-\sigma)\epsilon_{ij}$ captures consumer heterogeneity, and implies a correlation structure in the consumer-specific errors among subsets of goods. The random shocks ϵ_{ij} are assumed to be IID across consumers and products, and to follow the Type-I Extreme Value distribution. The random shock ζ_{ig}, in contrast, appears in consumer i's utility from all goods that belong in nest g. Its presence, therefore, induces correlation in the consumer's unobserved tastes within the nest. The distribution of $\zeta_{ig}(\sigma)$ depends on the parameter $\sigma \in [0, 1)$, and is shown by Cardell (1997) to be the unique distribution that induces $\zeta_{ig} + (1 - \sigma)\epsilon_{ij}$ to also have a Type-I Extreme Value distribution. This feature gives rise to simple analytical expressions for choice probabilities, facilitating the estimation and application of the model.

Substitution patterns. The key parameter governing substitution patterns in the nested logit model is σ. At one extreme, if $\sigma = 0$, the unobserved utility term reduces to the IID shocks ϵ, implying zero correlation in unobserved tastes within the nest. Such a pattern is consistent with McFadden's (1974) conditional logit model (see below), and implies that substitution away from any given product and towards other products is proportional to products' market shares. With a positive σ, unobserved tastes within the nest are correlated, suggesting stronger substitution towards products within the nest relative to products outside the nest. Since shocks are uncorrelated across nests, substitution towards products outside the focal product's nest is still proportional to these goods' market shares.[18] At the other extreme, as $\sigma \to 1$, products within the nest become stronger substitutes from the consumer's point of view. Estimating the correlation parameter σ can therefore provide valuable information regarding the degree of market segmentation.

Multiple-level models. The correlation structure above can be augmented by adding correlation levels. In principle, the sets $g = 1, ..., G$ can be repeatedly subdivided into subsets within which unobserved tastes are allowed to be correlated. The popular two-level nested logit (hereafter TLNL) is the most commonly-encountered case (Verboven 1996, Taylor 2014). Let us subdivide each set $g = 1, ..., G$ into mutually-exclusive subsets indexed by h_g. We then alter the utility function as follows:

$$u_{ij} = x_j\beta + \alpha p_j + \xi_j + \zeta_{ig} + (1 - \sigma_g)\zeta_{ih_g} + (1 - \sigma_g)(1 - \sigma_{h_g})\epsilon_{ij} \tag{2}$$

The model, therefore, has two correlation parameters, σ_g and σ_{h_g}, that capture correlation within the nest and within the sub-nest, respectively. Both parameters must belong in the interval $[0, 1)$. This model admits several interesting special cases. When $\sigma_g = 0$, the upper nesting level is irrelevant and correlation is only allowed within the sub-nests. When $\sigma_{h_g} = 0$, we essentially return to the one-level nested logit model presented above. When both parameters equal zero, no correlations are allowed and the model reverts to the simple conditional logit model.

Clearly, the choice of the nesting structure is of paramount importance for the empirical task discussed in this paper. The empirical application presented in Section 4 applies the TLNL to the Israeli dairy-based spreads market. There, the upper nesting level captures taste correlation within product category (e.g., Cottage cheese) while the lower nesting level captures correlation within category-brand combinations. In Section 3 we consider possible sources of guidance regarding the chosen structure, and discuss the manners with which the data allow us to reject particular structures, and to choose among competing structures if several such structures are considered.

Discussion: alternative demand models. While our focus in this paper is on the nested

[18] Substitution towards products within the focal product's nest is also proportional to these goods' market shares.

logit model, the literature offers a variety of discrete choice models. To name a few examples, Shaked and Sutton (1982) offer a theoretical vertical differentiation model, while Hausman and Wise (1978) estimate a conditional Probit model. While such models may be well-suited for particular applications, the vast majority of papers that estimate discrete choice models rely on models of the "logit" family, in which the nested logit belongs. This can be explained by the flexibility and computational advantages afforded by such models. Our brief discussion here would focus on comparing the nested logit model to other members of the logit family: the conditional logit model (McFadden 1974) and the Random Coefficient Logit model (BLP).

McFadden's (1974) model is equivalent to the one presented above with the parameter σ set to the value of zero. This model does not allow any correlation in consumers' unobserved tastes, and therefore produces unreasonable substitution patterns: substitution is proportional to the market shares of the products toward which it occurs (Berry 1994, BLP). According to this model, an increase in the price of a small, fuel-efficient automobile would not prompt consumers to predominantly substitute toward similar automobiles. Rather, the strongest flow of consumers would be toward the product with the largest market share, which may be a large automobile with low fuel efficiency. This model is, therefore, especially ill-suited for our purpose: defining competition groups.

BLP introduce the Random Coefficient Logit (RCL) model that accommodates a rich correlation structure, thus overcoming the limitations of the simple logit model. The RCL allows the utility weights β and α to be random coefficients, meaning that they are heterogeneous across consumers, and randomly drawn from a distribution that depends on estimated parameters. This model allows some consumers to place a higher-than-average weight on fuel efficiency. These consumers would tend to purchase fuel-efficient cars, and substitute toward other fuel-efficient cars when the price of the car of their choice increases. As a consequence, a substantial fraction of the substitution would be towards similar products, consistent with reasonable substitution patterns.

The Nested Logit is, in a sense, a compromise between the simple logit and the RCL. While it does allow for correlation structures, they are not as flexible as those of the RCL: correlation is only allowed within nests that are *a-priori* determined by the researcher, whereas the RCL allows these patterns to have a very rich structure, and for this structure to be determined by the data (i.e., by the estimated parameters that govern the distribution of the random coefficients). Verboven and Grigolon (2014) discuss the relative performance of the RCL versus the nested logit model (and also propose a model that combines both, the Random Coefficient Nested Logit). They find that, in the European car market, the RCL leads to wider market definitions relative to the nested logit, but that predicted price effects from mergers are very similar across the models. One should note that the expanding use of the RCL by practitioners, and the increasing availability of estimation methods and codes suggest an increased attractiveness of this model relative to the nested logit discussed in this paper.

Our focus on the nested logit model relies on two justifications: first, as discussed in the Section below, this model is amenable to empirical application using simple, linear estimation techniques, unlike the RCL which requires nonlinear estimation techniques involving costly simulation. In the context of policy applications in which practitioners need to provide a timely analysis, this poses a big advantage. A second justification is provided by the fact that applying multiple levels in the nested logit model (as suggested by the two-level model above) allows one to consider market segmentations that are increasingly rich and detailed.

3 Practical Implementation: Defining Competition Groups

This section describes the practical application of the model to data (with a specific case study presented in Section 4). Section 3.1 describes the estimation of the TLNL model, and Section 3.2 discusses the manners with which the data may reject the model. Section 3.3 describes the implementation of the SSNIP test and addresses specific challenges that arise in this task. Finally, Section 3.4 discusses practical aspects of constructing ψ-*Competition Groups*.

3.1 Estimating the TLNL Model

Following Berry (1994), estimating discrete choice models from aggregate market-level data is executed by inverting a "market share equation" that posits the equality of model-predicted choice probabilities to empirical market shares. The RCL, nested logit and simple logit models are all special cases of this approach. In the case of the simple logit and nested logit, one derives a linear estimation equation. In this paper we simply provide the estimation equation, referring the reader to Berry (1994) for further details on its derivation. The single-level nested logit model (equation (1), adding back the omitted time indices t) yields the following estimation equation, pertaining to product j (belonging in nest g) and market t:

$$ln(s_{jt}) - ln(s_{0t}) = x_{jt}\beta + \alpha p_{jt} + \sigma ln(s_{j/gt}) + \xi_{jt} \tag{3}$$

Equation (3) describes a linear regression model. An observation is a product-market combination j, t. The LHS is a function of the observed market shares for product j in market t, s_{jt}, and of the observed share of the outside option in that market, s_{0t}. The right hand side features the unobserved utility shifter ξ_{jt} as the econometric error term, as well as the following regressors: the vector of product characteristics x_{jt}, the price p_{jt}, and the term $ln(s_{j/gt})$, the natural log of the share of product j as a fraction of the total share of nest g in market t.

Two of the terms of the RHS of equation (3) are endogenous: p_{jt} and $ln(s_{j/gt})$, both of which are likely to be correlated with ξ_{jt}. Since firms observe this utility shifter when setting prices, it is correlated with p_{jt}. This shifter is also clearly correlated with market shares, including the share within the nest $s_{j/gt}$.

This endogeneity motivates estimating this linear model using instrumental variables (Two Stage Least Squares). The choice of instruments has been discussed extensively in the literature, and so would be discussed here only briefly. Cost shifters are natural instruments since they shift supply relations and prices while potentially being uncorrelated with the demand unobservable ξ_{jt} (Bresnahan 1989, Berry (1994)). Other typical instruments for price include functions of the characteristics x_{mt} of other products $m \neq j$. The characteristics of these products affect equilibrium prices.[19] We denote the vector of instrumental variables by z_{jt}.

The two-level model can similarly be used to derive a linear estimation equation. Defining $(1-\eta) = (1-\sigma_g)(1-\sigma_{h_g})$, the equation takes the following form (Verboven 1996, Taylor 2014):

$$ln(s_{jt}) - ln(s_{0t}) = x_{jt}\beta + \alpha p_{jt} + \eta ln(s_{j/h_g t}) + \sigma_g ln(s_{h/gt}) + \xi_{jt} \quad (4)$$

where $s_{j/h_g t}$ and $s_{h/gt}$ are product j's share out of sub-nest h's total share, and sub-nest h's share as a fraction of nest g's total share, respectively. Both these terms, and the price, are endogenous, implying that at least three instrumental variables are now necessary. As the equation demonstrates, the correlation parameter σ_g is directly estimated, while σ_h can be easily calculated as a function of the estimated parameters σ_g and η (implying that its standard error can be obtained via the Delta method).

The 2SLS estimates $\hat{\beta}, \hat{\alpha}, \sigma_g$ and η (or, in the case of the one-level model, $\hat{\beta}, \hat{\alpha}, \sigma$) are subsequently used to define *Relevant Markets* and *Competition Groups*, as explained in Sections 3.3 and 3.4 below. Before addressing these tasks, we next take up the question: how can the data inform us regarding the validity of the demand model specification?

3.2 Testability and Criteria for Model Selection

By using the nested logit model, the researcher effectively imposes an assumption regarding market segmentation. For instance, a one-level model that groups automobiles into nests such as "compact," "Sport Utility Vehicles" etc., allows SUVs to be closer substitutes to one another than they are to products that are not SUVs. Several questions arise: first, how should one decide how to allocate goods into nests? Second, how can the data reject certain correlation structures, or provide an indication that other structures are more plausible?

Our view on this issue is strongly motivated by the application at hand: defining markets and competition groups for antitrust purposes. We view the econometric procedures described in this paper as complementary to a qualitative analysis. Such an analysis involves, for instance,

[19] The identifying assumption is that these characteristics are uncorrelated with the utility errors. This assumption is valid if the characteristics are predetermined to the realization of these errors. The empirical literature offers a number of recent contributions that aim at relaxing such assumptions (see Crawford (2012) for a survey).

interviews with firms, analysts and consumers. The practitioner conducting these interviews may then form an opinion regarding the "true" market segmentation, or, at least, be able to a-priori rule out most possible structures and remain with a limited list of, say, two or three plausible structures. These structures depend on both the number of prominent product characteristics as well as on their order of importance in determining the resolution of market segmentation. For example, in the automobile market, the researcher might consider a one-level model based on the car type (compact, SUV, etc.), and a two-level model based on the car type as well as the gear type (manual, automatic, etc.). Translating each such structure into a nested logit model, one can use the estimated parameters in order to formally test such models, and to compare their performance.

In other words, we assume here that the task for the empirical application is not to provide a black box that magically produces the market segmentation that dominates all other possible structures. Rather, the starting point is a small set of alternative hypotheses regarding market segmentation. We first show how to test each such alternative model. If multiple examined models are not rejected, we suggest additional criteria that may help us choose among the surviving specifications.

3.2.1 Testing the Nested Logit model

Considering a specific correlation structure (i.e., a partition of the products $j = 1, ..., J$ into nests and possibly sub-nests), the structure should be considered *rejected by the data* if the estimated parameter values violate restrictions that are imposed by the theory. In our case, these formal restrictions are those that require the correlation parameters σ, σ_g and σ_h (depending on whether the one-level or two-level model is considered) to lie in the interval $[0, 1)$. Estimates that imply that such parameters are lower than zero or higher than 1 at a significance level of at least 95% should therefore be interpreted as a rejection of the specified model.

Some practical complications arise in this context. Specifically, the tested correlation structure is actually a family of models: the set of all models that share the same nesting structure, but differ in terms of the elements that enter the vectors x_{jt} and z_{jt}. That is, there may be many different sets of product characteristics and instrumental variables that may be utilized. Intuitively, the problem of multiple hypothesis testing (Holm 1979) is then relevant: if we test many such model variants, we would be likely to find at least some variants that would be rejected. One possibility is to use the Holm-Bonferroni correction and formally account for this issue. Alternatively, it may be practical to examine several specifications that make economic sense. If these are consistently rejected except for a small number of cases, one may deduce that the correlation structure is fundamentally rejected by the data. In contrast, if only a handful of specifications fail, one may be inclined to believe that the structure is consistent with the data. While the latter approach is informal, it may provide a reasonable rule of thumb: our trust in a correlation structure should decrease if it is often rejected by the data in a series of tests that consider different sets of utility shifters and instrumental variables. This principle would be demonstrated in the application of the Israeli dairy-based spreads market in Section 4.

Another important issue is that there may be some cases in which the model is not formally rejected, and yet there are reasons for the researcher to be skeptic about its ability to explain the data. One such case is estimated values of the σ correlation parameters that are very close to 1. Consider the one-level nested logit model, and suppose one obtains $\hat{\sigma} = 0.99$. From a theory perspective, no problem arises here: 0.99 is a valid value for the correlation parameter, and indicates near-perfect substitutability within the nest. In practice, however, one may suspect that this estimated value could reflect a failure to properly instrument for the endogeneity of the within-share variable $ln(s_{j/g})$. Notice that the left hand side variable includes the variable $ln(s_j)$, while on the right hand side we have $ln(s_{j/g}) = ln(s_j) - ln(s_g)$. In other words, applying OLS has the flavor of regressing $ln(s_j)$ on itself, which is likely to lead to an estimated parameter which is close to 1 in the finite sample. Properly instrumenting for $ln(s_{j/g})$ alleviates this mechanical relationship, and allows the finite-sample estimate of σ to move away from 1. None of the above, of course, suggests that an estimated value of 0.99 for σ is necessarily a consequence of misspecification. We only argue that such a value may indicate that a bigger effort is needed in constructing valid instrumental variables.

Other informal criteria to consider involve the economic content of the signs of the estimated values of β and α. As explained above, one should expect α to be negative. A long literature (Trajtenberg 1989, Berry 1994) discusses the fact that $\hat{\alpha}$ may be upward-biased as a consequence of the problem of price endogeneity. An estimated $\hat{\alpha} > 0$ may therefore, once again, be indicative of poor quality of instrumental variables, motivating a search for more efficient instruments. Counterintuitive utility coefficients (e.g., a negative utility weight on a desirable automobile feature such as the number of airbags) may similarly reflect specification errors, and should motivate the researcher to consider a wider set of specifications.

3.2.2 Comparing non-rejected correlation structures: elasticities and markups

Imagine that the researcher has a-priori narrowed down her set of plausible correlation structures to three, and that one of those structures has been rejected by the data in the spirit of subsection 3.1.1 above. This means that two different correlation structures (say, a one-level nested logit model and a two-level one) have not been rejected by the data. How should we choose among the two structures? We offer several possible answers to this question.

First, let us recall the objective of the analysis: defining *ψ-Competition Groups*. Given this stated goal, one may be inclined to compute these groups twice, using the two non-rejected correlation structures, one at a time (following the steps described below in subsections 3.3 and 3.4). If the result is qualitatively very similar, one may be less concerned with the question of which correlation structure provides a better description of preferences. In fact, one may then consider the conclusion regarding the ψ-*Competition Groups* more credible, as it is robust to the choice among the two demand models.

Notwithstanding the above possibility, in general it will not be true that the two structures yield the same conclusions regarding market segmentation, motivating us to think of ways of comparing the two models and choosing among them. Two natural avenues of addressing this task can be considered, and they both involve comparing some economic implications of the models—namely, elasticities and markups—to benchmark values that are considered by the researcher to be reasonable.

Demand elasticities. We note that the TLNL model implies the following probability that a random consumer chooses good $j \in g$ in market t:

$$Pr(j|x_t, p_t; \theta) = exp(\delta_{jt}/(1-\eta)) \cdot D_{ht}^{-\sigma_{h_{gt}}} \cdot D_{gt}^{-\sigma_{gt}} \cdot D_t^{-1} \tag{5}$$

where the following notation is used: the vectors x_t, p_t contain the observed product characteristics and prices of all goods offered in market t. The parameter vector θ contains all the model's parameters, i.e., $\theta = (\beta', \alpha, \sigma_g, \sigma_{h_g})'$. The term δ_{jt} is defined by $\delta_{jt} = x_{jt}\beta + \alpha p_{jt} + \xi_{jt}$. Finally, the following definitions apply:

$$D_{ht} = \sum_{k \in h_{gt}} exp(\delta_{kt}/(1-\eta)), \quad D_{gt} = \sum_{k \in gt} D_{kt}^{1-\sigma_{h_{gt}}}, \quad D_t = \sum_{g=0}^{G} D_{gt}^{1-\sigma_g}$$

Assuming that the number of consumers is large enough, this theoretical choice probability must be equal to the observed market share s_{jt}. As a consequence, one can analytically derive the market share derivative $\partial s_{jt}/\partial p_{jt}$ and therefore the own-price elasticity:

$$\frac{\partial s_{jt}}{\partial p_{jt}} \cdot \frac{p_{jt}}{s_{jt}} = \frac{\alpha}{1-\eta} p_{jt} \left[1 - \sigma_{h_g} s_{j/hgt} - \sigma_g (1-\sigma_{h_g}) s_{j/gt} - (1-\eta) s_{jt}\right] \tag{6}$$

By substituting in estimated parameter values and observed shares and prices, one can estimate the own-price elasticity for every sample product. Additional elasticities can be computed: as we discuss in more detail below, it is easy to obtain estimates of cross-price elasticities, arc elasticities (e.g. the percentage market share response to a 10% price increase), and the elasticity of demand toward a group of products (e.g., the response of the combined market share of all compact trucks to a 10% rise in the price of all compact trucks).

Such estimated elasticities can provide a "sanity check:" the estimates can be compared to relevant benchmarks familiar from the literature. An example is provided in Section 4, where the elasticities produced by certain estimated models are compared to a range of demand elasticities, for the same type of products considered, reported in a survey of many independent studies. Furthermore, obtaining reasonable demand elasticities can serve as a selection criteria among models that were not rejected by the data in the sense of section 3.2.1 above: suppose that the own-price elasticity of demand for a certain product is typically found in the literature to be in the range $[-2, -4]$. Suppose in addition that our two competing model variants provide estimated elasticities of (-3) and (-10), respectively. One may then view the first model as superior to the second in terms of capturing reasonable demand elasticities, and this criterion can be used to select among the two models.

Markups. Other economic implications of the estimated coefficients are markups. Following BLP and Nevo (2001), such an analysis begins with an assumption regarding firms' strategic behavior. The most common assumption is that of a differentiated-Bertrand price competition: firms simultaneously set prices to maximize profits. Omitting market indices t and denoting by $\mathcal{F}_f$ the set of products offered by firm f, this firm's profit function becomes:

$$\Pi(p_f, p_{-f}) = \sum_{j \in \mathcal{F}_f} \left[p_j \cdot M \cdot s_j(p_f, p_{-f}) - C(M \cdot s_j(p_f, p_{-f})) \right] \tag{7}$$

where the term p_{-f} is the vector prices charged by all competing firms. The first term inside the brackets is product j's revenue (price times quantity sold, which equals the market size M times the market share s_j). The second term is the cost of producing $M \cdot s_j$ units of good j. For simplicity, while this profit function can admit economies of scale in the production of a specific product, its separability in costs across products rules out economies of scope (those, however, can also be admitted in principle).

An interior, pure strategy Nash equilibrium in this game implies a system of equations comprised by J first order conditions, one for each product sold in the market.[20] Assuming in addition, again for the sake of tractability and simplicity, that the marginal cost c_j is constant in output, delivers the following system of equations in vector form:

$$p - c = \left(\Omega^* * S\right)^{-1} s(p) \tag{8}$$

where Ω^* is a $J \times J$ "ownership matrix" with a typical element $\Omega^*_{j,k}$ being equal to 1 if goods j and k are sold by the same firm, and equal to zero otherwise. This matrix is, of course, assumed to be observed data. The matrix S captures market share derivatives, such that $S_{jk} = -\partial s_k / \partial p_j$. The vectors p, c and s capture prices, marginal costs and market shares for all J products.

We refer to BLP and Nevo (2001) for additional details about this approach. For our current purposes, a key observation is that p, s and Ω^* are observed, while the matrix S can be computed given our estimated model (recall our above discussion regarding the computation of share derivatives and elasticities). It is, therefore, straightforward to compute the RHS of equation (8)

[20] The uniqueness of such an equilibrium is shown by Caplin and Nalebuff (1991) under strong assumptions such as a simple logit demand system and single-product firms. Nocke and Schutz (2015) provide an extention to the case of multi-product firms.

and obtain a predicted vector of markups $p - c$ (or, equivalently, a predicted value for marginal costs c, given that p is observed). This vector, in turn, can be used in a similar vein to how we used elasticities above: it provides both a "sanity check" and a possible criterion for comparison across candidate demand models (noting again that different estimated demand models would yield different estimated S matrices and hence different markup predictions). Practitioners (and certainly competition authorities) can often obtain some rough measure of costs and markups against which the predicted quantities can be compared. If the practitioner believes that a certain industry is characterized by thin margins, she would tend to favor a model that predicts low margins over a competing model that prescribes very high markups. Björnerstedt and Verboven (2013) provide an example: they use markups to choose among competing demand models in the Swedish Pharmaceutical market.

Nonetheless, the use of markups as a selection criterion should be done selectively and with caution. The main reason is that this criterion is only valid under the assumption made regarding the equilibrium. The markups computed above are only valid if a Nash-Bertrand equilibrium provides a reasonable approximation for firms' actual conduct.[21] There may be markets where conduct may differ considerably from that stated via a simple model: for example, in the Israeli dairy-based spreads market case study presented below, product prices were historically low due to price ceilings set by the government. Following a removal of many of these restrictions, a gradual process of price increases took place during the studied period. It is difficult to characterize this behavior as stemming from a consistent and stable equilibrium behavior. In contrast, the comparison of elasticities that we described above has the advantage of being robust to any assumptions regarding firms' supply-side behavior.

To summarize, our recommendation is to engage in the following multi-step process: first, the researcher should focus attention on a limited number of demand models that are consistent with quantitative and qualitative institutional details. Second, each such model should be estimated and tested, rejecting models that produce correlation parameters that are significantly smaller than zero or greater than 1. Using estimates from each model that survives the above steps, demand elasticities should be computed and compared to familiar benchmarks from the literature, followed by an elimination of models that produce unreasonable elasticities. Finally, surviving models can also be evaluated and compared based on their ability to produce realistic markups.

3.3 Implementation of the HMT Using the Estimated Demand System

Definition 1 stated that a set of products constitutes product j's Relevant Market if it is the smallest set that satisfies the SSNIP test. Therefore, in order to adequately delineate RM_j the

[21] Nevo (2001) shows how to compute markups under alternative assumptions, such as perfect collusion or single-product pricing. Björnerstedt and Verboven (2013) use a tuning parameter in the ownership matrix to consider conducts that are, informally, "intermediate" between perfect collusion, Nash-Bertrand and perfect competition assumptions.

formulation of two methodologies is required: how to verify that a proposed set satisfies the SSNIP test; and how to define the substitution ranking $\succ_j$ in order to confirm that it is indeed the smallest set that satisfies the test.
According to the AHMG, Section 4.1.1:

> "When applying the hypothetical monopolist test to define a market around a product offered by one of the merging firms, if the market includes a second product, the Agencies will normally also include a third product if that third product is a closer substitute for the first product than is the second product. The third product is a closer substitute if, in response to a SSNIP on the first product, greater revenues are diverted to the third product than to the second product."

If follows that the construction of the different sets that are candidates to be RM_j is not random and is in fact based on the substitution ranking. Thus, we begin by formulating this ranking before turning to characterize the practical methodology of the SSNIP test.

3.3.1 Substitution Ranking

According to the AHMG, product a is a closer substitute to the focal product than product b if given a SSNIP it holds that $E_{a,j} \cdot s_a > E_{b,j} \cdot s_b$, where $E_{k,j}$ is the arc price cross-elasticity of demand between product $k \in \{a, b\}$ and the focal product j. Using the midpoint arc-elasticity formula, $E_{k,j}$ is given by: $\frac{s_2^k - s_1^k}{(s_1^k + s_2^k)/2} / \frac{p_2^j - p_1^j}{(p_1^j + p_2^j)/2}$, where the index 1 represents the observed values of prices and market shares and the index 2 represents those values after a SSNIP in product j. The latter is obtained by applying the estimated parameters and the new price level of product j into δ_{jt} and using equation (5) to compute the predicted market shares of all products, including that of product i.[22] A possible alternative is to define $\succ_j$ using the point cross-elasticity rather than the arc cross-elasticity. An expansion of equation (6) to all possible cross-elasticities is given by:

$$\frac{\partial s_k}{\partial p_{jt}} \cdot \frac{p_{jt}}{s_{kt}} = \frac{\alpha}{1-\eta} p_{jt} \big[X_{jk}^1 - \sigma_{h_g} s_{j/hgt} X_{jk}^2 - \sigma_g (1 - \sigma_{h_g}) s_{j/gt} X_{jk}^3 - (1-\eta) s_{jt} \big] \tag{9}$$

In the base case where $X_{jk}^1 = X_{jk}^2 = X_{jk}^3 = 0$ the aforementioned equation depicts the cross-elasticity between two products of different nests. The case where both products are of the same nest ($j, k \in g$) is obtained by changing $X_{jk}^3 = 1$ and maintaining $X_{jk}^2 = X_{jk}^1 = 0$. If both products are of the same sub-nest ($j, k \in h_g$), then the corresponding cross-elasticity requires that $X_{jk}^2 = X_{jk}^3 = 1$ and $X_{jk}^1 = 0$. Finally, $X_{jk}^1 = X_{jk}^2 = X_{jk}^3 = 1$ yields the own-price elasticity.

[22] The updated consumer utility from purchasing product j is given by $\delta'_{jt} = \alpha(1 + SSNIP)p_{jt} + \beta x_{jt} + \xi_{jt}$. This change affects the predicted market shares of other products through the terms D_{ht}, D_{gt} and D_t. Note that in this case all other prices remain unchanged. A practical method to calculate δ'_{jt} is to use equation (4) and deduce that $\delta'_{jt} = ln(s_{jt}) - ln(s_{0t}) - \hat{\eta} ln(s_{j/h_g t}) - \hat{\sigma_g} ln(s_{h/gt}) + \hat{\alpha} \cdot SSNIP \cdot p_{jt}$.

If $E_{k,j}$ is the point cross-elasticity, then using $E_{k,j} \cdot s_{kt}$ as the measurement by which the order $\succ_j$ is determined is equivalent to using only the partial derivative $\partial s_{kt}/\partial p_{jt}$ since p_{jt} is a common factor to all products. Thus, allowing the order $\succ_j$ to be calculated directly from equation (9).

Substitution patterns. Since using the point cross-elasticity allows the use of a closed formula to calculate $\succ_j$, it also allows us to asses the effect the estimated correlation parameters (σ_g, σ_{h_g}) have on $\succ_j$. In general, the higher σ_g is, the higher the probability that RM_j will include only products from nest g such that $j \in g$.

3.3.2 Hypothetical Monopoly Test - Practical Methodology

After establishing the substitution ranking, the sets of products that are candidates to be the focal product's relevant market using the SSNIP test are the focal product alone, the focal product and its closest substitute, the focal product and its two closest substitutes, etc. We next discuss both the AHMG and the KS versions of the test, and then apply a refinement to the KS version to produce the form described in our formal model and used in the featured application in Section 4.

Following the AHMG, the selected set is the smallest one such that when a single firm owns all of its products, the firm's profit-maximizing price vector includes at least one product whose price increased by 5% relative to the observed price level.

Vertical structures, product ownership and profits. The existence of vertical chains suggests that the distinction between upstream and downstream profits may be important for performing the SSNIP. One possible strategy is to treat the retail margin as an additional cost to manufacturers in the spirit of Nevo (2001). Villas-Boas (2007) shows how to extend the first-order condition approach outlined above to infer both upstream and downstream margins given an estimated demand system. Finally, direct data on retail and producer margins may be collected by the researcher.

Another issue is ownership of other products. The focal product is produced by a certain manufacturer who may be producing other products as well. Performing the HMT disconnects the ownership of the focal product, and other products in the candidate set, from its original manufacturer and creates a new, virtual competitor in the market. Therefore, the profit-maximizing price vector of the hypothetical monopoly could be lower than the observed one as a result of this "new" competition, leading to broader market definitions. On the other hand, considering the focal product, as well as other products in the candidate set, to be part of an existing manufacturer's product portfolio is inconsistent with the concept of the HMT. As a result, the antitrust literature adopted a slightly different test, as noted in KS:

> "As a good working approximation, the profit-maximizing price increase is half as large as the maximal price increase that yields profits above their pre-merger level.

> Therefore the test should identify the narrowest group of products for which a 10% price increase of at least one of the products by a hypothetical monopolist would not result in a decrease in total profits."

Increase one price or all? The aforementioned test description asserts that the price of at least one product should increase for a set of products to be considered a relevant market for some focal product j. However, it does not specify which one or even if the focal product must be one of those products. The antitrust literature focuses on two end cases: increasing the price of the focal product only, or increasing the prices of all products in the candidate set. As discussed in Daljord et al. (2008), if all products share similar demand patterns and margins, then a symmetric price increase in all products is economically sensible. However, if certain products enjoy a distinctively high (or low) demand, then an asymmetric price increase should be preferred. Also, Coate and Fischer (2008) refer to the case in which all prices are increased as a special case that is designed to address situations of substantial sales at the "spatial fringes of the market".[23]

Measuring profits. At the beginning of this part we addressed the question concerning which profits should be maximized and measured. The question at hand is how should those profits be measured. We mentioned earlier that antitrust authorities have the prerogative to obtain data directly from the firms, and profit margins are no exception. However, sometimes such data are not available in the desired resolution, are incomplete due to various technical reasons, or are biased, as noted by Baumol (1996). In such cases, certain concessions are necessary, as demonstrated in Section 4. An alternative is to calculate markups and variable profits based on the first order conditions using equation (8). The choice among such alternatives should depend on the specifics of the application (e.g., the degree to which cost data are flawed).

3.4 Validation of $\mathcal{S}$ as a Competition Group

After establishing the mechanism by which relevant markets are defined, we can focus on the main objective of this work and calculate the ψ level of the different competition groups. In general, these competition groups will be the nests or sub-nests defined by the qualitative analysis, but this need not be the case and the ψ level of any set of products can be calculated according to Definition 3.

Since high ψ levels are desirable as they are an indication that the set $\mathcal{S}$ can be considered a proper CG, the main question that arises is how should the regulator treat low ψ levels. Two aspects of this question call for consideration: first, what is the minimal ψ level the regulator finds sufficient in order for that set to be viewed as a proper CG, and if that bar is not met, how

[23] In the application described below, examination of the data revealed that different products enjoy different margins, therefore a price increase was applied only to the focal product when conducting the SSNIP test. The formulation of the test in our definitions in Section 2 was done accordingly.

to resolve the contradiction between the qualitative and quantitative analysis regarding market delineation.

Clearly there should be some minimal ψ level below which the suggested segmentation cannot be considered valid for antitrust purposes. However, we do not presume to set such bar in this work and believe that it should be achieved through the working experience of competition authorities and court decisions. If the regulator does in fact find the measured ψ level to be low, this should motivate examining alternative partitions. Finally, it should be noted that extremely high ψ levels may indicate that the borders of the market were drawn too broadly. If the qualitative analysis was not uncontested, then narrower markets segmentations can also be considered, based on such an analysis.

4 Application: The Israeli Dairy-based Spreads Market

In this section we describe the results of the economic analysis conducted by the IAA of the Israeli dairy market as an example of the implementation of the concepts developed above.[24] We begin with some basic facts regarding the Israeli dairy market and the different players in it. Subsection 4.2 then briefly describes the data used in the analysis by the IAA. Subsection 4.3 is dedicated to the estimation equation and its components, mainly to the various instrumental and utility variables used. Finally, subsection 4.4 depicts the results of the HMT and the emanating competition groups.

4.1 Background

Dairy products are one of the largest food segments in Israel. Consumer expenditures on dairy products in 2015 exceeded 10 billion NIS (2.6 BN USD) which are approximately 11.2% percent of total food expenditure.[25] The Israeli dairy market is (almost) completely regulated. The Ministery of Agriculture determines raw milk production quotas and oversees their distribution across cowsheds. Moreover, raw milk price (the price dairies pay to farmers) as well the retail and consumer price of several dairy products (e.g., milk, white cheese, butter) is regulated as well, but not subsidized.

There are more than 90 active dairies in Israel. However, public estimates assert that the top 3 dairies process 95% of the local raw milk. Imports are equivalent to 15% of local production (in raw milk terms) and consist mainly of hard and semi-hard cheeses on account of their relatively

[24]This section does not contain and is not based on any data or analysis not previously disclosed in the IAA report on "Defining Markets Using Econometric Models of Demand", which the authors took part in drafting. The full report of the IAA with additional information and details can be found, in Hebrew, at http://www.antitrust.gov.il/subject/195/item/34421.aspx

[25]According to the Israeli Dairy Board, see http://www.halavi.org.il/info/idb/publications/PAGES1_80HIGH_NEW_WEB.pdf

long shelf life. This is due to high tariffs, limited import quotas, and high transportation and Kashruth costs. The distribution of market share between the top 3 dairies is as follows: Tnuva, which is the leading food supplier in Israel, is the largest dairy with a (monetary) market share of 55.1%; The Strauss Group (hereafter: "Strauss"), which is the second largest food supplier, has a market share of 21.5%; and Tara, which is a part of the CBC Group, the fourth largest food supplier, holds a market share of 7.8%.[26]

4.2 Data

Data sources. The data used in the IAA analysis come from two sources. First, retail sales data were obtained from Storenext database.[27] The data cover the Israeli dairy market over a period of of six and a half years (January 2005 to July 2011) and include monthly data on the quantity of units sold and the total amount paid by consumers (in ILS, including VAT) by SKU.[28] The identifying parameters of each SKU in the data include: category (e.g., white cheese), retailer (e.g.,Shufersal), format of sale (e.g.,heavy discount), manufacturer (e.g.,Tnuva), Weight (grams/ml), fat percentage (e.g., 5%) and flavor (e.g., regular). These data are used to define a product as a unique combination of observed characteristics.[29]

Second, production cost and wholesale data were obtained by the IAA from the three largest dairies in Israel. The same six and a half years are observed on a quarterly base (2005Q1 - 2011Q2, inclusive). Monthly production cost and wholesale data were derived using a common statistical technique.[30] These data sets were added to the previously described retail data set based on a categorial matching key.[31]

Spreads market. In discrete choice models the outside option should bear some economic sense of interchangeability with the examined products. For example, it is unlikely that a model meant to measure substitutability between different types of cars will include a category of cloths since the two categories do not fulfil the same need for the consumer. However, it is reasonable to consider one's need for mobility and include in the analysis public transportation or even bicycles. The Israeli dairy market is comprised of a wide range of products amongst which

[26] Dairy market shares are based on 2013 data, see http://www.globes.co.il/news/article.aspx?did=1000825042; food market shares are based on 2011 data, see http://www.globes.co.il/news/article.aspx?did=1000713569

[27] See http://www.storenext.co.il

[28] Due to the "cottage protest" outbreak at the end of June 2011 (Hendel et al. 2015) the last month was omitted from the data set.

[29] The four largest dairies in Israel account for more than 85% of sales (monetary values), therefore the data of all other suppliers were aggregated together and referred to as "Small".

[30] Data adjustment was carried out by the IAA for each category separately using a polynomial function of degree 6 based on the time variable. For each category c, note the cost or wholesale price data on month t by y_{ct} where $t \in \{1, .., 78\}$. Since there are 26 quarterly observations in each category, each three months belonging to the same quarter share the same monthly data. For example, for the first quarter in the sample $y_{c1} = y_{c2} = y_{c3} = y_c(2005Q1)$. The estimation equation for each category is then given by: $y_{ct} = \alpha + \beta_{c1}t + \beta_{c2}t^2 + ... + \beta_{c6}t^6 + \epsilon$. After estimating the coefficients, the expected cost or wholesale price, $\hat{y}_{ct}$ is calculated.

[31] In practice, the data cleansing process led to the use of the data of only one major dairy. Nevertheless, a positive correlation was found between all cost data as well as with several input indices.

various interchangeability relations might exist. Therefore, it is crucial to adequately define the consumer's purchase purpose. Three key questions were used by the IAA to discern between products: Who is the target consumer, when is the product consumed during the day, and how is the product consumed?[32] In light of these questions the IAA divided all dairy products to three parent groups: (i) a "spreads" group for products that can be consumed with a knife or eaten on a substrate (like bread); (ii) a "beverages" group for products that can be consumed in a glass; (iii) a "cups" group for products that can be consumed with a spoon (like yogurt). We consider this categorization to be in accordance with the Brown Shoe reference of a "broad market" in which distinctive sub-markets may exist. As the IAA's quantitative analysis focused solely on the spreads group and its composition, the term $\mathcal{J}_t$ represents the set of "spreads" products available to consumers on a certain month.

After defining the consumer's purchase purpose, an estimation of the potential spreads market size is required. Since a valid assessment regarding per capita consumption of spreads (including non-dairy spreads like jam, guacamole, etc.) was unavailable, the IAA examined three alternative scenarios for the potential monthly market size (in Kg) to assess the results' sensitivity using the following equation:[33]

$$M_t = population_t \times 0.9 \times Consumption \times \frac{365}{12}$$
$$Consumption = \{0.05, 0.125, 0.25\}$$

The examined daily consumption values induce an outside option market share ranging from 60% to 95%. The results presented in this section are based on a per capita daily consumption of 125 grams, generating a market share of 90% attributed to the outside option.[34]

Product characteristics. After establishing the broader purchase purpose and the potential market size, a formal definition of the different products meeting that purpose is required. One option is simply to address each SKU as a unique product. However, this seems to be impractical since the data consist of hundreds of different SKUs. Naturally, an aggregation of several SKUs together is in order. Table 1 describes the four key characteristics of each SKU that were used so that (almost) each combination uniquely defines a product.

Three technical notes: First, all of the products considered in this analysis are sold on the retailer's shelf in closed packages and not in customized sizes in delis. Second, product sales were summed across all available points of sale. Third, products' prices were normalized per Kg and package size did not qualify as a characteristic.[35]

[32] For example: chocolate milk is usually consumed by children, in the morning, without additional supplements.

[33] Population size was obtained by the IAA from Israeli CBS assuming a constant increase rate throughout the year. Multiplying by 0.9 is designed to exclude ages $0-4$.

[34] Sensitivity analysis, similar as in Villas-Boas (2007), indicate that market size does not have substantial effect on the eventual competition groups.

[35] Examination of the data revealed that in several categories there is a quantity discount such that SKUs sold in large

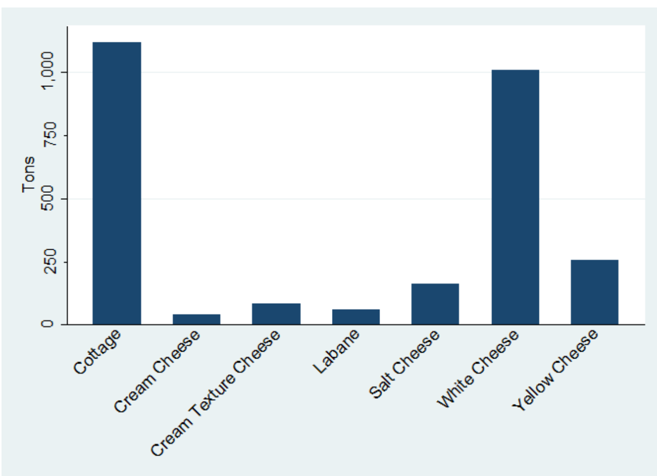
Table 1: Product Characteristics

Category	Manufacturer	Flavor	Fat Level
Cottage Cheese	Tnuva	Regular	
White Cheese	Strauss	Olives	
Yellow Cheese	Tara	Garlic	Varies across categories
Salt Cheese	Gad	Onion	
Cream Texture Cheese	Small	Other	
Cream Cheese			
Labane			

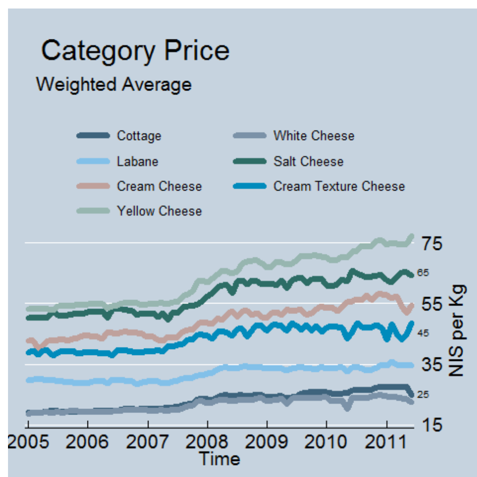
Yellow cheese is the common name for various hard and semi-hard cheeses like Edam, Gouda and Emmental; Salt cheese is a cheese whose NaCl contents are at least 2% of its weight; Cream texture cheese resembles cream cheese only with lower fat contents (usually 5% as opposed to over 15%); Labane is produced by draining whey out of yogurt, creating a sour flavored cheese.

Figure 1: Category Descriptives - Monthly Averages

(a) Quantity Sold



(b) Monthly Price



Descriptive statistics. Each observation in the final database is a unique product-month combination. It includes the product's total quantity and monetary sales; its cost of production and its four distinguishable characteristics. There are 151 different products sold over a 78 months period, creating $8,001$ unique observations (not all products were sold on each period). Figures 1a and 1b illustrate the monthly average quantity sold of each category and the evolution of prices throughout the sample period.

packages are cheaper per same weight unit. These discount rates were weighted by the relevant product's average price and were not perceived to be of value by themselves. This approach is consistent with the empirical literature, see Nevo (2001) and Ackerberg (2003).

4.3 Demand Estimation

Nest structure. The qualitative analysis held by the IAA indicated that the two most important parameters to consumers are the category of the product and its manufacturer. Accordingly, the IAA chose to apply a TLNL model in which the product's category determines the first level of the hierarchy and the product's manufacturer determines the second.[36] As was stated earlier, this structure can be rejected by the data either on a technical basis (invalid coefficients) or an economic reasonableness basis (e.g, extreme elasticities). Given this structure, economic significance can be attributed to each argument in the estimation equation, equation (4), which we restate here:

$$ln(s_{jt}) - ln(s_{0t}) = x_{jt}\beta + \alpha p_{jt} + \eta ln(s_{j/h_g t}) + \sigma_g ln(s_{h/gt}) + \xi_{jt}$$

In this equation g represents the product's category (e.g., cottage cheese); h represents the product's manufacturer (e.g., Tnuva); and j is the specific product (e.g., labane by Gad, 11% fat, regular flavor). The two market shares arguments are the manufacturer's market share in a specific category ($s_{h|g}$) and the product's share out of all of the manufacturer's products in a specific category ($s_{j|h_g}$).

Instrumental variables. As discussed above, the application of the TLNL requires at least three instrumental variables, one for each endogenous variable. We briefly review the variables that were used in the IAA's work and the endogenous variables they are correlated with.

Production cost - An increase in production cost is expected to have a similar effect on consumer price as firms usually transfer at least some of the additional costs to the retailer, who in turn increases the price to the end consumer. Thus, making this variable a relevant IV for the endogenous price variable.

Ownership - On January 2008 Apax Partners completed the transaction giving them control over 77% of Tnuva, the largest dairy in Israel. A dummy variable divides the data into two distinctive periods - prior to the change in ownership and after it. This exogenous change in the identity of decision makers is a relevant IV for the endogenous price variable.

Competition variables - Following Section 3.1, several variables that capture the competition intensity in each category were considered. Such variables are common in the empirical literature (see BLP and Verboven (1996)) and while various configurations were examined, the two presented here are:[37]

- The number of additional products of the manufacturer in a specific category - This variable

[36] The hypothesis that each category is considered to be a CG can also be tested using a one-level model. The more complex TLNL model was used as it may be more applicable in future studies. It should be noted that the hypothesis that categories constitute CGs is consistent with the predictions of either demand model.

[37] Verboven examines the automobile market and uses the average of competitors' observed product characteristics as an instrument variable while BLP use a more sophisticated function of competitor's characteristics. The IAA conducted sensitivity analysis using additional competition variables which did not lead to different results.

is expected to be positively correlated with the price as a wider variety of products allows the firm to raise prices since at least some of the churn will be to her other products. Since one objective of expanding product variety is increased market share, a positive correlation with the manufacturer's market share of the category ($s_{h|g}$) as well as a negative correlation with the product's market share ($s_{j|h_g}$) due to high in-house competition are expected.[38]

- The number of additional products of the manufacturer in a specific category within the same fat level - Each category was divided into three fat levels (low, medium, high) based on each product fat contents. Similar to the previous competition variable, this variable is expected to be positively correlated with the manufacturer's market share of the category, $s_{h|g}$. With respect to the product's market share out of all of the manufacturers products in a specific category, $s_{j|h_g}$, the correlation is ambiguous. On the one hand, a negative correlation can be expected due to more intense competition, while on the other hand, a proliferation of products in a specific fat-level may indicate there is high demand for such products.[39]

Consumer utility variables. Apart from the endogenous variables, the right hand side of the estimation equation includes the observed characteristics (x_j) as well as the unobserved (ξ_j). The x_j vector includes:

Hierarchy position - Dummy variables were assigned to indicate both the nest (category) and sub-nest (manufacturer) to which the product belongs to. In addition, an interaction variable of category and manufacturer was included to allow for a different manufacturer effect in a given category.

Product characteristics - Based on the four observed product characteristics the following variables were constructed to account for consumer utility:

- Flavor dummy variables.
- Weighted average fat percent (hereafter "Average fat").[40]
- Squared average fat - allowa for non-linear (e.g., decreasing marginal utility) effect on consumer utility.
- Time-average fat interaction - allows for changes in consumer preferences over time, such as health trends.
- Category-average fat interaction - Fat levels vary across categories, therefore it is necessary to allow for different effects on consumer utility based on the product's category.

[38]The data values range between 0 and 10.

[39]The data values range between 0 and 5.

[40]A product was defined as an aggregation of several SKUs with similar characteristics, one of which is the SKU's fat level. Some products were defined over a range of fat levels (e.g., yellow cheese, regular flavor of 28% fat or above) and as such include SKUs with various fat levels causing the product's fat level to vary over time. In practice most products were defined by a specific fat level, based on the dominant SKUs in that category.

Time variables - to account for periodic effects a time trend variable was included.[41] In addition, since dairy products consumption increases significantly on Shavuot, which is a Jewish holiday celebrated on either May or June, a dummy variable indicating the occurrence of the "Shavuot" holiday is used.

Results. Estimation results using 2SLS are reported next. The validity of the four instrumental variables can be assessed from their first stage coefficients:

Table 2: First Stage, IV Coefficients

First stage	p	$s_{j\|h_g}$	$s_{h\|g}$
Production cost	0.05535	0.01453	-0.02125
Ownership	2.9126***	-0.06243	0.04979
Additional products	0.9768***	−0.62827***	0.11891***
Additional products in fat level	−1.75441***	0.25913***	0.0506***
R-squared	0.7491	0.4087	0.8767

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Three of the instrumental variables are significant in at least one out of the three first stage regressions. Moreover, the effects of the competition variables are consistent with the predicted effects discussed above.[42] It should be noted that although a positive correlation between the cost and the price was expected, the effect is not significant.

Second stage endogenous variables' coefficient results are displayed in Table 3. Estimation results suggest that, as expected, consumer price has a negative effect on consumer utility. Also, the category level correlation (σ_g) and manufacturer level correlation (σ_{h_g}) are 0.811 and 0.578, respectively. Their presence in the range of $[0, 1)$ means that the suggested structure is not rejected by the data. Moreover, the fact that these coefficients are significantly higher than zero is consistent with the product segmentation portrayed by the two-level hierarchy.

Table 3: Second Stage, Endogenous Variables Coefficients

Second stage	Coefficient
α	-0.0156***
η	0.92043***
σ_g	0.81122***

*** $p < 0.001$.

To provide some basic intuition regrading these results, Table 4 provides inter and intra-nests median cross-price elasticities. Two main conclusions arise from these elasticities: first,

[41] Sensitivity analysis using time dummies for each period found no effect on market definition results.

[42] The positive correlation with $s_{j|h_g}$ can be interpreted as an indication of the consumers' preference of a certain fat level in each category.

consistent with the high level of σ_g, a change in one product's price bears virtually no effect on the quantity consumed of products from other categories, regardless of the manufacturer's identity. Second, due to the positive value of σ_{h_g}, the cross elasticity between products of the same manufacturer within category is higher than that of two products within the same category produced by different manufacturers. Furthermore, arc-elasticities were computed as described in Section 3 and were found to be consistent with the range of values reported in the meta-analysis of Andreyeva et al. (2010).

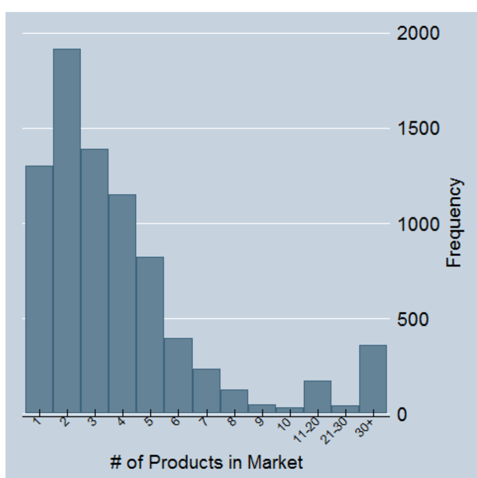
Table 4: Median Cross-price Elasticities

Product affiliation	Median elasticity
Intra-category, same manufacturer	0.12967
Intra-category, other manufacturers	0.03033
Inter-category, regardless of manufacturer	0.0001

4.4 The Hypothetical Monopolist and Market Segmentation Tests

The IAA applied the practical HMT mechanism described in section 3 to the Israeli dairy-based spreads market data, and generated the relevant market for each of the 8,001 product-period combinations. The distribution of relevant market sizes is illustrated in Figure 2.

Figure 2: Distribution of Market Definitions, 10% Price Increase



The figure shows that more than 90% of relevant markets include up to 7 additional products besides the focal product. However, approximately 5% of the relevant markets are comprised of more than 20 products, indicating that for certain products the relevant market spans several categories.[43]

Ψ-Competition groups. The qualitative analysis conducted by the IAA asserted that each category can be viewed as a CG. However, the existence of relevant markets that exceed the boundaries of one category indicates that for certain periods in the sample, some of the categories cannot be considered a CG. Using ψ-competition groups allows for measuring how limited is the market share of products whose market definition includes products outside of the CG. The larger the ψ values, the more inclined the researcher will be to accept the proposed partition of CGs. Table 5 describes the minimal ψ levels of each category across time, according to definition 3.[44]

Table 5: Minimal ψ Levels

Category	Minimal ψ
Cottage Cheese	100%
White Cheese	99.99%
Yellow Cheese	73.58%
Salt Cheese	97.94%
Cream Cheese	96.53%
Cream Texture Cheese	100%
Labane	99.4%

Consistently with the qualitative analysis of the IAA, in every category except for yellow cheese, the relevant market of the absolute majority of products remains within the boundaries of that category throughout the entire sample. The interesting question that arises from Table 5 is how should the ψ level of the yellow cheese category be interpreted. As discussed in Section 3, when ψ values are considered to be low, which was the case here from a conservative perspective, they indicate that the a-priori segmentation imposed by the researcher does not fully capture consumers' preferences. Therefore, the IAA applied a tweak to the hierarchy structure: the yellow cheese category was separated into two different categories- fat (over 9% fat) and low-fat (up to 9% fat).[45] After recalculating each product's relevant market under the new nest structure, the

[43] It may be the case that these broad market definitions are the result of the "cellophane effect". That is, the pre-existing high pricing of those products due to the exercising of market power caused the additional price increase to lead consumers to switch to other products, even though they were not close substitutes to begin with. Nevertheless, the total sales volume of these products is usually very small, even negligible, thus invalidating the concern that the suggested categories are not CGs. Also, some relevant markets include only the focal product and indicate an incentive to increase prices over the observed level. However, this "sub-optimal" pricing may not necessarily be a result of firms not maximizing profits for two main reasons: first, in the FMCG markets, and especially in dairy products markets, the price updating process is usually done gradually. It was previously shown that a price increase trend was evident in all of the considered categories. Therefore, it is not unlikely that in several periods some products were underpriced. Second, the HMT cannot account for the firms' inability to increase prices due to non-monetary considerations, such as public-opinion or fear of regulatory intervention (e.g., price regulation) as described in Hendel et al. (2015). As a result, the existence of single-object market definitions does not compromise the validity of the test.

[44] For clarification, subtracting the numbers in the table from 100 represents the maximal market share of products whose relevant market included at least one product out of the considered category, throughout the sample.

[45] In-depth examination of relevant markets of yellow cheese products found that most products whose relevant market

new minimal ψ levels were 97.43% for fat yellow cheese and 80.14% for low-fat yellow cheese.[46]

Sensitivity analyses. After establishing the model is not rejected by the data and that it generates predictions that are consistent with both the qualitative analysis and the existing literature, the IAA examined its robustness. Two different parent groups of sensitivity analyses were carried out: variables composition and data contents.

- Variables composition - tests that demonstrate that results were not obtained due to an esoteric or incidental variable added to the estimation equation. Changes to both IVs and utility variables were examined.
- Data contents - tests that examine changes with respect to the potential market size as well as to the calculation of firms' profit margins.

All of the analyses support the IAA's findings and are brought in greater detail in Appendix A. Furthermore, following the suggested method in Section 3.2 to discern between different hierarchy structures, the IAA examined several alternative hierarchies, and these results are also reported in Appendix A. These models were examined even though they were not suggested by the qualitative analysis to demonstrate that the data can reject other structures.

5 Concluding Remarks

This paper uses an empirical demand estimation model, the nested logit, as a tool for performing market definition. This work fills a gap in the literature by addressing the application of such demand estimates to the SSNIP test. We point out some practical difficulties in performing the SSNIP using an estimated demand system and suggest a refined version of the test based on the works of Katz and Shapiro (2002) and Daljord et al. (2008). Moreover, we demonstrate the practical implementation of the SSNIP test using the case-study of the Israeli dairy market.

Following Brown Shoe, we introduce a new concept of "competition groups" in an attempt to allow the researcher to examine whether a specific set of products demonstrates a high degree of substitutability and can therefore be considered as a well-defined market. The featured example demonstrates the use of a refined concept, the ψ-competition group, that calculates the fraction of the segment's products which market definition does not exceed the segment's borders. An important question for future research is how large should the market share of such "confined" products be in order for the entire set to be considered a proper CG.

exceeded the yellow cheese category were low-fat.

[46] ψ levels of other categories remained practically unchanged.

References

Daniel A Ackerberg. Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, 44(3):1007–1040, 2003.

Tatiana Andreyeva, Michael W Long, and Kelly D Brownell. The impact of food prices on consumption: A systematic review of research on the price elasticity of demand for food. *American Journal of Public Health*, 100(2):216–222, 2010.

Jonathan B Baker. Stepping out in an old brown shoe: In qualified praise of submarkets. *Antitrust Law Journal*, 68(1):203–218, 2000.

William J Baumol. Predation and the logic of the average variable cost test. *Journal of Law and Economics*, 39 (1):49–72, 1996.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 63(4):841–890, 1995.

Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262, 1994.

Jonas Björnerstedt and Frank Verboven. Does merger simulation work? a natural experiment in the swedish analgesics market, 2013.

Timothy F Bresnahan. Departures from marginal-cost pricing in the american automobile industry: Estimates for 1977–1978. *Journal of Econometrics*, 17(2):201–227, 1981.

Timothy F Bresnahan. Competition and collusion in the american automobile industry: The 1955 price war. *The Journal of Industrial Economics*, 35(4):457–482, 1987.

Timothy F Bresnahan. Empirical studies of industries with market power. In Schmalensee R. and Willig R., editors, *Handbook of Industrial Organization*, pages 1011–1057. Elsevier, 1989.

Andrew Caplin and Barry Nalebuff. Aggregation and imperfect competition: On the existence of equilibrium. *Econometrica: Journal of the Econometric Society*, 59(1):25–59, 1991.

N Scott Cardell. Salford systems. *Econometric Theory*, 13(2):185–213, 1997.

Malcolm B Coate and Jeffrey H Fischer. A practical guide to the hypothetical monopolist test for market definition. *Journal of Competition Law and Economics*, 4(4):1031–1063, 2008.

Gregory S Crawford. Endogenous product choice: A progress report. *International Journal of Industrial Organization*, 30(3):315–320, 2012.

Øystein Daljord. An exact arithmetic ssnip test for asymmetric products. *Journal of Competition Law and Economics*, 5(3):563–569, 2009.

Øystein Daljord and Lars Sørgard. Single-product versus uniform ssnips. *International Review of Law and Economics*, 31(2):142–146, 2011.

Øystein Daljord, Lars Sørgard, and Øyvind Thomassen. The ssnip test and market definition with the aggregate diversion ratio: A reply to katz and shapiro. *Journal of Competition Law and Economics*, 4(2):263–270, 2008.

Alon Eizenberg. Upstream innovation and product variety in the us home pc market. *The Review of Economic Studies*, 81(3):1003–1045, 2014.

Joseph Farrell and Carl Shapiro. Improving critical loss analysis. *Competition Policy Center*, 2007.

Barry C Harris and Joseph J Simons. Focusing market definition: How much substitution is necessary? *Research in Law and Economics*, 12:207–226, 1989.

Jerry A Hausman and David A Wise. A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica: Journal of the Econometric Society*, 46(2):403–426, 1978.

Igal Hendel, Saul Lach, and Yossi Spiegel. Consumers' activism: the facebook boycott of cottage cheese. 2015.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2): 65–70, 1979.

Kai Hüschelrath. Detection of anticompetitive horizontal mergers. *Journal of Competition Law and Economics*, 5(4):683–721, 2009.

Ulrich Kaiser. An empirical test of models explaining research expenditures and research cooperation: Evidence for the german service sector. *International Journal of Industrial Organization*, 20(6):747–774, 2002.

Michael L Katz and Carl Shapiro. Critical loss: Let's tell the whole story. *Antitrust*, 17(2):49–56, 2002.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In P Zarembka, editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, 1974.

Daniel McFadden et al. *Modelling the Choice of Residential Location.* Institute of Transportation Studies, University of California, 1978.

Serge Moresi, Steven C Salop, and John Woodbury. Implementing the hypothetical monopolist ssnip test with multi-product firms. *Available at SSRN 1087536*, 2008.

Aviv Nevo. A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics & Management Strategy*, 9(4):513–548, 2000.

Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342, 2001.

Volker Nocke and Nicolas Schutz. Multi-product firm oligopoly: An aggregative game approach. *mimeo*, 2015.

Daniel P O'Brien and Abraham L Wickelgren. A critical analysis of critical loss analysis. *Antitrust Law Journal*, 71(1):161–184, 2003.

David Pettit. Submarkets and supermarkets: Ftc v. whole foods market and the resurrection of brown shoe. *Geo. Mason L. Rev.*, 16:971–1011, 2008.

David Scheffman and JJ Simons. The state of critical loss analysis: Lets make sure we understand the whole story. *Antitrust Source*, 3(2):1–9, 2003.

Avner Shaked and John Sutton. Relaxing price competition through product differentiation. *The Review of Economic Studies*, 49(1):3–13, 1982.

Yair Taylor. *Essays on Innovation, Competition and Regulation in the Pharmaceutical Industry.* PhD thesis, Duke University, 2014.

Manuel Trajtenberg. The welfare analysis of product innovations, with an application to computed tomography scanners. *The Journal of Political Economy*, 97(2):444–479, 1989.

Frank Verboven. International price discrimination in the european car market. *The RAND Journal of Economics*, 27(2):240–268, 1996.

Frank Verboven and Laura Grigolon. ested nor random coefficients logit? a comparison between alternative discrete choice models of product differentiation. *Review of Economics and Statistics*, 96(5):916–935, 2014.

Sofia Berto Villas-Boas. Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies*, 74(2):625–652, 2007.

Gregory J Werden and Luke M Froeb. The effects of mergers in differentiated products industries: Logit demand and merger policy. *Journal of Law, Economics, & Organization*, 10(2):407–426, 1994.

A Sensitivity Analyses

Alternative hierarchies. As was mentioned in Section 3.2, the qualitative analysis may point at several different structures as the "true" market segmentation. Although the qualitative analysis of the IAA in the case of the Israeli dairy products market was straightforward, several alternative structures were examined strictly for the purpose of demonstrating different ways in which the data can reject "false" structures. The following are the results of several alternative structures, the hypotheses that underly them, whether they were rejected or not, and why:

- One-level nest, based on category - This structure corresponds with the hypothesis that each category is a CG that was tested in the featured model. No significant difference in either category own-price elasticities or HMT results was found.
- Two-level nests, opposite order (manufacturer as the first level and category as the second) - The resulting category own-price elasticities are 3-6 times higher relative to the baseline model. Furthermore, in 5 out of 7 categories own-price elasticities exceeded the maximal elasticity reported in Andreyeva et al. (2010). Thus, even though this structure is not formally rejected by the data it seems to be an inferior alternative.
- One-level nest, based on fat level - This structure corresponds with the hypothesis that each fat-level constitutes a CG, regardless of the division to categories or the manufacturer's identity. Under this structure the σ value exceeded 1 in several specifications.[47] In the few valid cases, category own-price elasticities were considerably above the reported elasticities in Andreyeva, et al., (2010).

Variables composition.

- Instrumental variables - Different compositions of competition variables were considered. Additional variables that were tested include the number of additional products in the category, the number of additional products in the category with the same fat level and the number of additional products in the category with the same flavor, of the same manufacturer. Some of the variables were indeed statistically significant, still market segmentation tests regarding all categories through most of the periods were not affected by instrumental variables composition.
- Utility variables - Since all observed product characteristics were included in the estimation equation, robustness tests narrow to various interactions and time effects. Changing the time trend variable into time dummy variables as well as adding an interaction between the time trend and the different categories did not affect HMT results significantly.

[47] Different compositions of both instrumental and utility variables that were modified to comply with the new structure were examined.

Data contents.

- Potential market size - The size of M directly affects the size of the outside option and therefore the dependent variable. Since a substantiated estimation of the potential market could not be found, two scenarios that constitute a decrease of 60% and a 100% increase with respect to the baseline model were examined. In both cases, changes to the potential size of the market did not affect the HMT results.
- Profitability margins - As we mentioned in Section 4.3, the IAA analysis uses data from only one of the largest dairies to calculate the profitability margins of each product while performing the HMT. Applying different margins to several other firms by 10% or 20% did not lead to significantly different results.