Some patterns of market shares of brands within and across product categories

Rajeev Kohli and Raaj Sah

University of Chicago, Harris School of Public Policy. Working paper series 06.04. December 2005.

ABSTRACT

This paper: (i) reports an empirical regularity in the market shares of brands; (ii) presents a theoretical framework for understanding the observed regularity; (iii) adduces additional empirical consequences of the framework, which are some counterintuitive relationships among market shares of brands across different product categories; and (iv) presents empirical evidence for these consequences, thus providing additional support for the theoretical framework. Our cross-sectional data on market shares consists of 1171 brands in 91 product categories of foods and sporting goods sold in the US. The key empirical regularity is that, in each category, the decrease in the market share between two successively ranked brands becomes smaller as one progresses from higher-ranked to lower-ranked brands. The power law represents these patterns well, in an absolute sense, and better than an alternative model, namely, the exponential form, which has been studied in the literature but without having been compared to any alternative. The latter form predicts that the ratio of the market shares of any two successively ranked brands is a constant. We present some potential implications of our findings. We also offer an interpretation of the previously known square-root relationship between market share and the order of entry of firms into an industry. The theoretical framework that we present for understanding the patterns reported here shares its foundation with that of the familiar Dirichlet-multinomial paradigm of brand purchases. This framework has some intuitive interpretations; it accommodates multiple product categories; and it allows for the entry and exit of brands over time.

Published version:

Kohli, Rajeev and Raaj Sah. “Some empirical regularities in market shares.” *Management Science*, Volume 52, Number 11, November 2006, pages 1792-1798.

HARRIS SCHOOL WORKING PAPER
SERIES 06.04

SOME PATTERNS OF MARKET SHARES OF BRANDS WITHIN AND ACROSS PRODUCT CATEGORIES

Rajeev Kohli and Raaj Sah

SOME PATTERNS OF MARKET SHARES OF BRANDS WITHIN AND ACROSS PRODUCT CATEGORIES

Rajeev Kohli
Columbia University

Raaj Sah
University of Chicago

December 6, 2005

"...not only can choice mimic chance, but chance can mimic choice."

Maurice G. Kendall,
Presidential Address, Royal Statistical Society

Parts of this paper are included in the article, "Some Empirical Regularities in Market Shares," forthcoming in *Management Science.*We are grateful to Asim Ansari for his help with the hierarchical Bayes estimates of random-coefficients models. We thank Petr Barton, Noel Capon, Sunil Gupta, Mac Hulbert, Kamel Jedidi, Don Lehmann, Robert Lucas, Puneet Manchanda, Natalie Mizik, Irina Ostrovnaya, Suchandra Thapa and Anne Thomas for their help. The paper has benefited from numerous comments and suggestions from seminar participants at Columbia University, Harvard University, Singapore Management University and the University of Pittsburgh.

SOME PATTERNS OF MARKET SHARES OF BRANDS WITHIN AND ACROSS PRODUCT CATEGORIES

Abstract

This paper: (i) reports an empirical regularity in the market shares of brands; (ii) presents a theoretical framework for understanding the observed regularity; (iii) adduces additional empirical consequences of the framework, which are some counterintuitive relationships among market shares of brands across different product categories; and (iv) presents empirical evidence for these consequences, thus providing additional support for the theoretical framework. Our cross-sectional data on market shares consists of 1171 brands in 91 product categories of foods and sporting goods sold in the US. The key empirical regularity is that, in each category, the decrease in the market share between two successively ranked brands becomes smaller as one progresses from higher-ranked to lower-ranked brands. The power law represents these patterns well, in an absolute sense, and better than an alternative model, namely, the exponential form, which has been studied in the literature but without having been compared to any alternative. The latter form predicts that the ratio of the market shares of any two successively ranked brands is a constant. We present some potential implications of our findings. We also offer an interpretation of the previously known square-root relationship between market share and the order of entry of firms into an industry. The theoretical framework that we present for understanding the patterns reported here shares its foundation with that of the familiar Dirichlet-multinomial paradigm of brand purchases. This framework has some intuitive interpretations; it accommodates multiple product categories; and it allows for the entry and exit of brands over time.

I. INTRODUCTION

Bass (1995, p. G7) defines an empirical generalization as "a pattern or regularity that repeats over different circumstances and that can be described by simple mathematical, graphic, or symbolic methods." Unlike a law of classical physics, an empirical generalization is not necessarily universal, and its parameters are not necessarily invariant across different circumstances. Empirical generalizations are typically approximate rather than exact, and they are descriptive rather than directly causal (Ehrenberg 1982). These generalizations facilitate the important task of constructing theories (Ehrenberg 1995) and of generating testable consequences beyond the original data (Simon 1968). The Pareto income distribution is an enduring empirical generalization in economics (Persky 1992). The Bass diffusion model (Bass 1969) and the Dirichlet-multinomial (DM) paradigm of brand purchases (Goodhardt, Ehrenberg and Chatfield 1984) are important empirical generalizations in marketing.

We present a number of empirical findings concerning patterns of market shares of brands. We also present some of the implications of these findings for marketing practice and research. Our data, described in detail later, consists of 506 brands in 48 product categories of foods and 665 brands in 43 product categories of sporting goods sold in the US. Examples of product categories of foods are orange juices and breakfast cereals. Examples of brands within the product category of orange juices are Minute Maid and Tropicana. We describe a theoretical framework for understanding these observed patterns. This framework has many appealing features for the study of market shares. It has intuitive implications concerning the growth rates of the market shares. Also, it shares its probabilistic foundation with that of the familiar DM paradigm of brand purchases noted above. We adduce empirical consequences of this theoretical framework, which are some counterintuitive relationships among market shares of brands across different product categories. We present empirical evidence for these consequences, thus providing additional support for the framework.

In the spirit noted in the first paragraph, our analysis is descriptive and approximate, rather than causal and exact. The power law is a central organizing concept of our analysis. In the context of brands, it says that if brand j has the r_j-th highest market share s_j, then $s_j = A(a + r_j)^{-b}$, where A, a, and b are constants. A core empirical content of the power law is that the decrease in the market share between two successively ranked brands becomes smaller as one progresses from higher-ranked to lower-ranked brands. A testable alternative to the power law, which we examine, is that the ratio of market shares of any two successively ranked brands is a constant. This implies the exponential form, $s_j = Ge^{-gr_j}$, where G and g are constants.

Our main empirical findings, when product categories are considered one at a time, are as follows: (1) The power law holds very well in an absolute sense; the R^2 values are consistently over 0.90 and often over 0.95. (2) The power law describes the data significantly better than the exponential form. (3) The relative superiority of the power law over the exponential form is greater for those product categories that have lower values of b. This is consistent with Mandelbrot's (1963) theoretical prediction that a power law with sufficiently large values of the coefficient b approximates the exponential form. (4) The exponential form fits better for lower-ranking market shares than for higher-ranking market shares.[1]

We also present findings from some "derived" data sets on market shares. As will be seen later, they contribute to a theoretical understanding of the power law. The three derived data sets, described in detail later, are created by: (a) pooling the market shares across all product categories; (b) considering those brands that occupy the highest rank in their respective product categories; and (c) taking the averages of the market shares of brands that hold a particular rank in their respective product categories, then considering these averages across the ranks. We show that all of the four results stated in the previous paragraph hold for each of the three derived data sets just described. These regularities are counterintuitive in the sense that there are no *a priori* reasons to expect that any such patterns exist among the market shares of brands across product categories.

We have abstracted from several important topics related to the present paper, in order to keep the paper within reasonable bounds of length and scope. Among these are: (i) potential implications of our findings for government policies, including those towards market dominance; (ii) the possibility that analogues of our results from the derived data sets may exist for phenomena other than market shares, such as the distribution of city sizes or of individuals' incomes; and (iii) general issues which arise from the fact that the power law describes many patterns in the human, physical and biological worlds (Brock (1999) and Gell-Mann (1994) provide perspectives and citations). At the end of this paper, in the concluding section, we describe many of the ways in which our empirical analysis can be extended with data which are more comprehensive than those which we have used. We also discuss there some of the ways in which the present descriptive and macroscopic analysis is complementary to causal microscopic analysis of the market shares of brands in any given product category.

Contributions of the present paper. To our knowledge, this paper is the first to examine the power law for a large number of product categories, rather than for just one or another product category. These findings suggest, in the spirit described at the beginning of the paper, that the power law is a candidate to be considered as an empirical regularity for the market shares of brands. Quite apart from this, there are potential disadvantages of examining the power law while limiting oneself to just one or another product category. For example, as will be seen later, the presence of the

[1] In an important study on market shares, Buzzell (1981) tested the exponential form, without explicitly comparing it to an alternative specification. Like him, we find that the exponential form fits the data reasonably well. However, our findings, including those summarized in this paragraph, suggest that the power law is a better description of the data than the exponential form. Separately, our analysis does not support two other hypotheses on market shares: (i) the "rule of three and four" of the Boston Consulting Group (1987), which predicts that the three largest market shares will be in the ratio 4:2:1; and (ii) Kotler's hypothesis (1977) that the top three brands will have 40%, 30% and 20% market shares.

power law may then be attributed to the special characteristics of those categories, without perhaps recognizing its widespread prevalence. In the literature, examples of the application of the power law to individual product categories are Chung and Cox (1994) on the number of hit albums produced by a music group, Adamic and Huberman (2000) on the number of links pointing to a web site, and Kalyanaram, Robinson and Urban (1995) on a relationship between market share and the order of entry of firms producing prescription anti-ulcer drugs and certain packaged consumer goods. Later, as a part of the discussion of the implications of our findings, we discuss some differences between our findings and those of the last study just mentioned.

To our knowledge, this paper is also the first to report a set of empirical patterns in the derived data sets. Our theoretical framework, which is not limited to one or another product category, provides a possible way to understand the findings from these data sets.

Organization of the paper. Section II presents some preliminaries. Section III describes the data and the estimation procedures. Section IV summarizes the empirical findings. Section V presents a theoretical framework. Section VI summarizes some of the implications of our findings for marketing practice and research. The concluding section contains brief remarks on some of the important research topics which are related to or follow from the present paper.

II. SOME PRELIMINARIES

Consider one product category with n brands; we will later deal with multiple product categories. Suppose, for now, that no two brands have the same market shares; we will deal later with ties in market shares. We assign the index $j = 1, 2, 3, \ldots, n$ to brands in decreasing order of market shares. Let s_j denote the market share of brand j. Let r_j denote the market-share rank associated with brand j. Since there are no ties here, $r_j = j$, but, as will be seen later, this need not always be the case. The power law and the exponential form are respectively:[2]

$$s_j = A(a + r_j)^{-b}; \text{ and} \tag{1}$$

$$s_j = Ge^{-gr_j}. \tag{2}$$

Empirical contents of the power law and the exponential form. Define the "share ratio" of two successively ranked brands as $f_j \equiv s_j/s_{j+1}$. We assume that $A > 0, a > -1$, and $b > 0$ for the power law, and that $G > 0$ and $g > 0$ for the exponential

[2] Sometimes the power law is expressed as $s_j = A'(a' + hr_j)^{-b}$, where A', a', h and b are constants. This yields (1) by setting $A \equiv A'h^{-b}$ and $a \equiv a'/h$. Expression (1) or its special cases are often referred to, without consistency, as the Pareto Law, the discrete Pareto distribution, and Zipf's Law. A widely used special case is that with $a = 0$; the framework in Section V refers to this version. The special case with $a = 0$ and $b = 1$ has been used extensively in the study of city sizes; see Gabaix (1999). Kalyanaram, Robinson and Urban (1995), cited earlier, use the special case with $a = 0$ and $b = 1/2$.

form. These inequalities ensure that $s_j > 0$ and $f_j > 1$. The share ratio is $f_j = [1 + (1/(a + r_j))]^b$ for the power law and $f_j = e^g$ for the exponential form. It follows that $f_j > f_{j+1}$ for the power law and $f_j = f_{j+1}$ for the exponential form. As noted earlier, this is a crucial difference between the empirical contents of (1) and (2). The scaling coefficients A and G play no role in these empirical contents. We illustrate this for A; the arguments for G are analogous. If we consider all of the brands in a product category, then from (1) and $\sum_j s_j = 1$, we obtain $A = 1/\sum_j (a+r_j)^{-b}$. If the brands below some level of market share are excluded from consideration (for practical reasons, data sets on market shares typically exclude brands with very small market shares), then from (1) and $\sum_j s_j < 1$, we obtain $A < 1/\sum_j (a + r_j)^{-b}$. In either case, the scaling coefficient A does not affect the empirical content of the power law noted earlier. Further, each market share is less than unity because, as discussed earlier, the market share for each brand is positive, and because $\sum_j s_j = 1$ or $\sum_j s_j < 1$, depending on whether we consider all brands in the product category or whether the brands below some level of market share are excluded.[3]

III. DATA AND EMPIRICAL METHODS

We examine two sets of data. The first, made available by Nielsen Market Research, reports the market shares of 506 brands in 48 product categories of foods. These market shares are for a large urban market in the Southwestern US, aggregated over the 120 weeks from January 1993 to May 1995. The second data set is published by the Sporting Goods Association of America. It contains the market shares in the US for 665 brands in 43 product categories of sporting goods, aggregated over the 1999 calendar year. The first column in Table 1 displays the number of brands in each product category for foods. The names of the product categories are not displayed in this table because these were withheld by those providing the data. Table 2 lists the names of the product categories for sporting goods and the number of brands in each product category. In these tables, we have displayed the product categories in descending order of the number of brands within a category. This mode of presentation of product categories will be helpful later.

As is the case with most data sets for market shares of brands, these two data sets reflect the motivations and constraints of those who created them. In brief, the following aspects seem noteworthy: (a) The data on foods, collected at the store level for a smaller geographical area, is perhaps more accurate than that on sporting goods. A limitation of the former data is the exclusion of certain types of stores from the Nielsen audits. (b) The construction of product categories is not based on explicit considerations of empirical research. For example, we do not have a random

[3]The notion of a "long tail," sometimes associated with the power law, can be understood in the present context as follows. If the number of brands is large then, from $f_j > f_{j+1}$, many brands at the lower end of ranks will have market shares which are quite comparable to one another. Such a long tail will not arise if the number of brands is small. The empirical content of the power law is orthogonal to the presence or absence of a long tail.

selection of product categories of foods and sporting goods. We have used all of the data available to us. (c) All market shares are in equivalent (quantity) units, without distinguishing brand variations and SKUs. For example, Minute Maid orange juice is sold in different variations, such as with or without pulp or calcium, and in various sizes. The data on the sales of the Minute Maid brand add the units (that is, gallons) across variations and sizes. (d) The data exclude brands with market shares smaller than 1%. Such exclusions of small brands are, for practical reasons, common among data sets on market shares. (e) As seen in Tables 1 and 2, the number of brands is small for several product categories. This aspect is also common among data sets on market shares. It arises in part because of the exclusions just noted and in part because some product categories are dominated by a few large brands.

For the above reasons, and also because we have examined a total of only 91 product categories, the patterns reported in this paper are tentative, and it is an open question whether such patterns exist in other data sets. At the same time, many of these caveats are partly ameliorated by our findings. Our two data sets represent a relatively broad range of products in their respective markets, namely, for foods and sporting goods. These two markets are quite unrelated, including regarding consumers' reasons for buying or not buying particular goods or brands, and producers' methods of selling their products. The two data sets differ in the length of time over which the data has been collected; one year for foods as against 18 months for sporting goods. The data sets also differ in their geographical coverage; regional for foods as opposed to national for sporting goods. Moreover, these two data sets have been constructed by two different organizations under different procedures and with different objectives, without any coordination with each other. Notwithstanding these key differences, our empirical findings from the two sets of data are very similar. This could be viewed as a partial indication that our results and conclusions are likely to be robust in spite of the unique characteristics or limitations of the data sets.

Estimation methods. One set of estimates presented in this paper consists of the parameters of the power law and the exponential form for each of the product categories of foods and sporting goods. For brevity, we refer to these as "category-specific" parameters. This shorthand also reduces the possibility of confusion between these parameters, which are specific to each product category, and those that are estimated for each of the three derived data sets, which, it may be recalled, contain data selected across product categories. All of our parameter estimates for the derived data sets are based on the minimization of the sum of squared-errors (MSSE).

We employ two different estimation methods for category-specific parameters. The first method is the MSSE. Under this method, category-specific parameters are estimated using only the data for the product category under consideration. The second method is hierarchical Bayes estimation of random-coefficients models (RCM). Under this method, we estimate category-specific parameters by pooling the data for all product categories. Each of these two methods has its advantages and disadvantages. Among the advantages of the MSSE is that it is simple and familiar, and its estimates are unbiased. At the same time, as will be seen below, these estimates have certain limitations, especially for the product categories that have a very small number of brands. An advantage of the RCM is that in some sense it makes better use of the available data. Another advantage is that it allows for the comparison of non-nested models, taking into account the fact that the exponential form has one less parameter than the power law (Kass and Raftery 1995). A caveat in the use of the RCM is that, as stated earlier, the product categories in our data are not a random sample of the respective universes of the product categories of foods or sporting goods. Also, the RCM estimates are biased, because these are posteriors obtained by combining common priors with category-specific information. What is noteworthy from our point of view is that the MSSE and the RCM are mutually complementary, in that they yield the same qualitative conclusions. We describe the details of the MSSE below in the main body of the paper, and those of the RCM in the appendix. As seen in the appendix, the RCM estimates show an overwhelming superiority of the power law over the exponential form. The values of the Bayes factors, which are commonly used for comparing such non-nested models, are e^{263} for the product categories of foods and e^{423} for those of sporting goods.

In certain cases, the market shares in a product category are identical up to two decimal places. In such cases, we assign the same average rank to these tied data points. Thus, for instance, if the market shares are identical for the two brands below the highest-ranked brand, then $r_2 = r_3 = 2.5$.

We use a nonlinear procedure to estimate the parameters of the power law by rewriting (1) as $\ln s_j = \ln A - b \ln(a + r_j)$. For the exponential form, we linearly estimate the parameters by rewriting (2) as $\ln s_j = \ln G - g r_j$.

As we noted earlier, the number of brands is small for several product categories. For example, as seen in the lower rows of Table 1, there are 6 or fewer brands in each of 16 out of 48 product categories of foods. This aspect of the data has some consequences for the MSSE estimates of category-specific parameters.[4] One consequence is that, as is to be expected, the parameter estimates based on too few data points will likely have large standard errors. This typically turns out to be the case, especially for some of the estimates of the power law. Another consequence is as follows. For some product categories, the R^2 values for the power law estimates increase monotonically with the value of b, and this increase is nearly imperceptible for values of b larger than 10. For example, an increase in the value of b from 10 to 50 typically increases the R^2 value by less than 0.01, on base values of R^2 generally in excess of 0.9. These product categories have two distinguishing features, which can be seen in Tables 1 and 2, in comparison to those in which this issue concerning the value of b does not arise. First, these product categories typically, but not in every case, have fewer brands. Second, the R^2 values for the power law for these product categories are very close to the corresponding R^2 values for the exponential form. This latter

[4]These consequences do not arise for the RCM estimates. Recall that, under this method, category-specific parameter estimates are obtained by pooling data for all product categories.

feature suggests an interpretation in view of Mandelbrot's prediction, cited earlier, that a power law with sufficiently large values of b approximates an exponential form. The interpretation is that, within the limitations of the data, the power law and the exponential form are roughly equally good descriptions for these product categories, and, thus, besides being large, the value of b does not indicate anything in addition to the preceding interpretation. Keeping this in mind, for these product categories, we have used a value of 10 for b in the parameter estimates presented in the paper; estimates with larger values of b are available upon request from the authors.

Constraints on the magnitudes of market shares. The constraints on these magnitudes are: (i) that a market shares should be positive and less than unity; and (ii) the adding-up condition, that the sum of the market shares should be no larger than unity, given that our data sets exclude brands with less than 1% market share. In Section II, we showed that these constraints are satisfied by the theoretical statements of the power law and the exponential form. One empirical approach with regard to these constraints is to build them within the estimation procedures, so that the estimates of the parameters will, by design, satisfy these constraints. Another approach, which we follow here for the reasons that will become apparent below, is to obtain unconstrained estimates of the parameters, and then to assess whether the estimates are consistent with one or another of such constraints.[5] We present below our findings in this regard for the power law; those for the exponential form are analogous.

For all of our estimates, we find that the predicted market share of each brand is positive and less than unity. We also find that the adding-up condition is satisfied for each of an overwhelming proportion of our product categories (84 out of 91). A reason for the seven exceptions is as follows. The estimated market shares, and therefore their sum, are random variables because the estimates of the parameters are random variables. The sum of the predicted market shares is therefore more likely to exceed unity for a product category for which the sum of the actual market shares is closer to unity. Among the exceptions, the smallest sum of the actual market shares is 0.968 and the largest is 0.997. For these exceptions, the quantitative conclusions concerning the power law remain unchanged if the adding-up condition is built into the estimations; these estimates are available from the authors.

[5] This approach has the advantage that it potentially illustrates additional strengths of the theoretical specifications and empirical procedures. Analogous issues have been debated for decades in the literature in economics on the estimation of systems of equations describing consumers' expenditures on various categories of goods and services; see Deaton and Muellbauer (1980, Chapter 3) for a review of this literature which began in the 1930s. The neoclassical economic theory of consumer demand suggests several restrictions on the parameters of such equations, in addition to the restriction that the "budget shares" (that is, the shares of the expenditure on various categories) should add up to one. The early research in this area generally tended to incorporate such restrictions into the estimation procedures, so that the estimates will tautologically satisfy these restrictions. The subsequent literature (for example, Christensen, Jorgenson and Lau 1975) has generally been in favor of obtaining the parameter estimates without bringing such restrictions into the picture, and then analyzing the extent to which the estimations satisfy such restrictions.

IV. EMPIRICAL RESULTS

Table 1 displays the estimated parameters, and the corresponding values of R^2, for the power law and the exponential form, separately for each of the 48 food categories. Table 2 displays the corresponding results for the 43 sporting-goods categories. These results suggest that the power law holds well in an absolute sense. For example, for foods, the value of R^2 for the power law is greater than or equal to 0.95 for 37 out of a total of 48 product categories, and it is larger than 0.9 for 44 product categories. For sporting goods, the value of R^2 for the power law is greater than or equal to 0.95 for 41 out of a total of 43 product categories, and it is larger than 0.9 for all 43 product categories. A value of one is displayed for the R^2 in some cases in this paper because we have rounded off these values to two places after the decimal.

Parameter estimates (of a, g and unrestricted b) that are statistically significant at the 95% confidence level are shown in boldface in Tables 1 and 2. Error estimates are not applicable if the reported value of b is 10, given the restriction mentioned earlier. The overall picture in this regard is that, if a parameter estimate is not significant, then it is typically but not always the case that the corresponding product category has a small number of brands. This can be seen in two different ways in Tables 1 and 2, in which, as was noted earlier, the product categories are displayed in descending order of the number of brands within a category. First, compared to food categories, many more of the estimates of b are significant among sporting goods categories, which typically also have more brands within individual categories than food categories. Second, the estimates of b that are significant are concentrated more among the upper rows of each of these two tables than among the lower rows.

Note in Tables 1 and 2 that the estimated parameters are different for different product categories. This is what we would expect. Suppose, to the contrary, that the parameters were the same for two product categories which have the same number of brands. Then the brands with the same rank will have identical market shares across these two categories. This, on a systematic basis, will be almost entirely contrary to what we know about market shares.

Figure 1 shows how the values of R^2 for the power law and the exponential form differ at different values of b.[6] The upper panel in Figure 1 is for food categories. The values of b for food product categories are taken from Table 1, and the product categories are reordered according to ascending values of b. The numbers displayed on the horizontal axis of this panel are the labels of the product categories, after this reordering. These numbers are in themselves not relevant to what this figure shows. The vertical axis displays the corresponding values of R^2 for the power law as well as for the exponential form. The lower panel of Figure 1 presents the corresponding

[6] The R^2 value here refers to the proportion of explained variance in the logarithm of the values of market shares for each product category. As we use a non-linear estimation procedure for the power-law model, the proportion of explained variance does not have the usual statistical properties associated with R^2 in a linear regression model. However, it is still a reasonable measure for the limited purpose of comparing the fits of the two models.

results for sporting goods. These two panels show that, for lower values of b, the values of R^2 for the power law are substantially larger than those for the exponential form, and that the values of R^2 from the two models are less distinguishable at higher values of b. These findings are consistent with the theoretical prediction of Mandelbrot (1963) that a power law with large values of b approximates the exponential form.

We now set up the apparatus to describe, graphically and mathematically, the derived data sets that we create by separating and combining, in particular ways, the raw data across product categories. All of the derived data sets are industry specific; that is, the data for foods are not combined with the data for sporting goods. Figure 2 is a graphical aid to understanding the derived data sets; the upper panel is for foods and the lower panel is for sporting goods. This figure displays the market shares for each rank across product categories. That is, the highest-ranked market shares for various product categories form the vertical cluster at rank = 1, the second-highest-ranked shares form the vertical cluster at rank = 2, and so on. Note that the neighboring clusters in Figure 2 overlap in their vertical ranges. However, their means are quite well separated, as will be seen later in Figure 5.

We now introduce some notation to deal with multiple product categories and to formally define the derived data sets. The index $i = 1$ to m represents product categories. Thus, m is 48 for foods. Within product category i, there are n_i brands. We assign the index $j = 1$ to n_i to the brands in non-increasing order of market shares, breaking ties arbitrarily. After the ties are broken using the protocol described earlier, we assign the rank r_{ij} to brand j in category i. The market share of brand j in product category i is s_{ij}. Let $H \equiv \max_i n_i$ denote the largest number of brands in any of the m product categories. Thus, as Table 1 shows, $H = 27$ is the largest number of brands in any food category. For each $1 \leq h \leq H$, we define a set $\psi(h)$ whose elements are the indices of product categories with h or more brands; that is, $\psi(h) = \{i | h \leq n_i, 1 \leq i \leq m\}$. Then the data in Figure 2 is described as follows. The vertical cluster at rank = 1 displays the numbers $\{s_{i1} | i \in \psi(1)\}$, the vertical cluster at rank = 2 displays the numbers $\{s_{i2} | i \in \psi(2)\}$, and so on.

The first derived data set, presented in Figure 3, pools the market shares of brands across all product categories. Put differently, this figure displays the observations contained in all of the vertical clusters in Figure 2, combined together, and then rearranged in descending order. Formally, Figure 3 displays $\{s_{ij} | 1 \leq j \leq n_i, 1 \leq i \leq m\}$, rearranged in descending order. The upper two panels are for foods and the lower two panels are for sporting goods. The left panels are for the power law and the right panels are for the exponential form. In addition to the data, each panel presents the parameter estimates (namely, a and b for the power law, and g for the exponential form), the value of R^2, and a line that describes the pattern predicted by the estimated parameters. We follow the same conventions for graphical presentations in later figures.

The second derived data set, presented in Figure 4, contains the market shares of brands that hold the highest rank in their respective product categories. Thus, this figure displays the numbers that form the vertical cluster at rank = 1 in Figure 2, ordered by their rank within this cluster. Formally, these numbers are $\{s_{i1}|i \in \psi(1)\}$, rearranged in descending order.

The third derived data set, presented in Figure 5, displays the averages of the market shares of brands that hold a particular rank in their respective product categories. For example, the market share displayed at rank = 1 in Figure 5 is the average of the vertical cluster at rank = 1 in Figure 2. Formally, Figure 5 displays $\{\sum_{i\in\psi(h)} s_{ih}/|\psi(h)|,\ 1 \leq h \leq H\}$, rearranged in descending order, where $|\psi(h)|$ is the number of elements in $\psi(h)$.

As can be seen from Figures 3, 4 and 5, the power law holds very well in an absolute sense. All of the estimates (of a, g and unrestricted b) presented in these figures are significant at the 5% level. In Figures 3 and 5, the values of R^2 for the power law are larger than those for the exponential form. In Figure 4, the value of R^2 for the power law is comparable to that for the exponential form. Mandelbrot's prediction, discussed earlier, applies in this case.

Finally, recall our conclusion that the exponential form fits better for lower-ranking market shares than for higher-ranking market shares. Consider Figure 3 as an illustration; analogous observations hold for all of our other results. The lines in the right-hand panels of this figure describe the pattern predicted by the estimated parameters of the exponential form. The fits provided by these lines are markedly better for lower-ranking market shares than for higher-ranking market shares. There is no such visual asymmetry in the fits provided by the power law, which are presented in the left-hand panels of Figure 3.

V. A THEORETICAL FRAMEWORK

In this section, we describe and motivate a theoretical framework that relates to our empirical findings and to several other observations presented earlier. This framework is based on a model of Hill (1970, 1974) and its subsequent developments (Hill and Woodroofe 1975, Chen 1978); see Aoki (1996, pp. 226–236) for a partial but succinct summary. Among the strengths of this framework are that it leads to some intuitive interpretations; it connects with more than one part of marketing literature which have been developed independently of issues of interest here; it accommodates multiple product categories; and it allows for the entry and exit of brands. We begin this section with a brief description of the DM distribution and some of its special cases. This will facilitate the rest of this section because this distribution is the foundation of our theoretical framework. We then note those aspects of Hill's model that are pertinent for the present paper. This is followed by a discussion of some predictions of Hill's model in relation to our empirical findings. We then describe some of the ways in which our theoretical framework is connected to marketing literature. This is followed by a brief discussion of an alternative theoretical model of the power law. At the end of this section, we present some caveats pertaining to our theoretical framework.

The DM distribution. For expositional simplicity, we begin with only one product category that has n brands. Define $p \equiv (p_1, p_2, p_3, \ldots, p_n)$, where p_j denotes the probability that a consumer buys one unit of brand j. Define $\ell \equiv (\ell_1, \ell_2, \ell_3, \ldots, \ell_n)$, where ℓ_j is the number of units of brand j purchased by the consumer, and define $N \equiv \sum \ell_j$. For brevity, in the preceding expression and in the rest of this section, we suppress the range of the index j over sums and products; j ranges from 1 to n. The purchase of brands is a zero-order process. Given p, the probability that the consumer buys ℓ_j units of brand j is given by the multinomial distribution:

$$\Pr(\ell|p) = N! \prod p_j^{\ell_j}/\ell_j!. \tag{3}$$

The heterogeneity in consumer preferences is represented by the specification that p has a Dirichlet distribution with parameters $(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n)$. That is,

$$\Pr(p) = \frac{\Gamma(\sum \alpha_j)}{\prod \Gamma(\alpha_j)} \prod p_j^{\alpha_j - 1}. \tag{4}$$

We assume throughout that each α_j is positive. Expressions (3) and (4) yield the following unconditional probability that consumers buy ℓ_j units of each brand j:

$$\Pr(\ell) = N! \frac{\Gamma(\sum \alpha_i)}{\Gamma(N + \sum \alpha_i)} \cdot \frac{\prod \Gamma(\ell_j + \alpha_j)}{\prod \{\Gamma(\alpha_j)\Gamma(\ell_j + 1)\}}, \tag{5}$$

which is the DM distribution. If $\alpha_j = \alpha$, for all $j = 1, \ldots, n$, then the distribution (4) of the heterogeneity in preferences is symmetric. If $\alpha = 1$, then the DM distribution (5) reduces to Bose–Einstein statistics (Fader and Schmittlein 1993, p. 481)

$$\Pr(\ell) = 1 \Big/ \binom{N-1}{n-1}. \tag{6}$$

If $\alpha = \infty$, then (5) reduces to the Maxwell–Boltzmann distribution.[7] Some other special cases are discussed later.

Hill's model. A derivation of Hill's model and its subsequent developments will be redundant and far too detailed for our purpose. Here we note some of its critical elements, assumptions, and predictions. For vividness, this abstract model uses the language of species, genera, and families. In its lieu, given our context, we use the language of market shares, brands, product categories, and so on. In Hill's results, a sufficient condition concerning the consumers' behavior is that $\alpha = 1$, and he conjectures (Hill 1974, p. 1024) that his results hold for all finite values α. With a sufficient condition that α is finite, Chen (1978) proves many parts of this conjecture, including for the first result discussed in the next subsection, and, to our knowledge, the remaining parts have not so far been refuted. It is further possible that the

[7]See Feller (1968, pp. 38–43) for a discussion of Bose–Einstein statistics and the Maxwell–Boltzmann distribution.

predictions described in the next subsection do not necessarily require the assumption that (4) be symmetric. That is, these predictions may hold without any restrictions on the values of $(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n)$, except that each of these parameters is finite in value.

Hill's model incorporates multiple product categories and allows the number of brands in each category to be a random variable, thereby accommodating the introduction and withdrawal of brands from the market for various product categories. Let m denote the number of product categories, where m is large. The product category i has n_i brands with total sales N_i. Leaving aside some technical conditions, Hill's model assumes that: (i) purchases are independent across product categories; (ii) the number n_i of brands within a category is random; and (iii) n_i/N_i is independent across product categories. This model uses asymptotic arguments. Our heuristic numerical simulations suggest that the patterns predicted by the model become increasingly recognizable as the value of one or more n_i is increased, and that the patterns become reasonably recognizable if $n_i > 10$ for all i.

Predictions. One prediction concerning the derived data sets is that the power law arises, in an approximate sense, when the market shares are pooled across product categories. As shown in Figure 3, our findings support this prediction. Another prediction is that the power law arises among the highest-ranked market shares across product categories. As shown in Figure 4, and as discussed earlier, the power law fits this derived data set well. Finally, we have presented a pattern in Figure 5 which is not predicted by Hill's model. Recall that, in this figure, we: (i) calculated the mean of the k-th largest market shares across product categories, and (ii) arranged these means in descending order. Our analysis suggests that the power law is a reasonable description of this derived data set.

Relationships between the theoretical framework and marketing literature. The following are among the relationships: (i) Expression (5) for the DM distribution has been widely studied and validated in marketing literature for analyzing brand purchases of consumers (Goodhardt, Ehrenberg and Chatfield 1984). (ii) Recall that the parameters $(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n)$ represent the heterogeneity in consumer preferences, and that the distribution (4) of this heterogeneity is symmetric if $\alpha_j = \alpha$, for all $j = 1, \ldots, n$. For a simple interpretation of symmetric preferences, consider its special case in which $n = 2$. Then (4) becomes a Beta density with parameters (α, α). If $\alpha = 1$, this Beta density implies a uniform distribution of preferences for brands within the product category. The heterogeneity distribution is U-shaped if $\alpha < 1$, representing a symmetric polarization of preferences for brands. If $\alpha > 1$, the heterogeneity distribution is unimodal, and its variance around the mode decreases as the value of α increases. (iii) If consumers buy sufficiently large quantities of each brand, then an intuitive implication of the DM distribution (5) is that the expected probability that the next unit purchased will be of a particular brand is proportional to the number of units of this brand already purchased. This implication follows an expression derived by Fader and Schmittlein (1993, p. 481). (iv) The preceding implication is also obtained from Bose–Einstein statistics without the need to make any assumption concerning the magnitudes of previous purchases by consumers. This conclusion follows from the expression just cited.

An alternative model of the power law. The power law has been derived from Gibrat's Law (Ijiri and Simon 1975), which is well known. In our context, this law posits that period-to-period changes in the sales of a brand follow a process in which the probability of any specified percentage change is independent of the brand's present size. In other words, the sales of each brand have the same probability of increasing or decreasing by 5 percent, 10 percent, or any other fixed amount, regardless of current sales. Note that this implication is the same in its content as those noted in parts (iii) and (iv) of the previous subsection. This is not a coincidence. Ijiri and Simon (1975) have shown that Bose–Einstein statistics can be derived from Gibrat's Law.

We have used the DM distribution as the starting point of our theoretical framework and not Gibrat's Law, even though the preceding discussion suggests that the two share some similarities in their implications. One reason for this choice of ours is that, as seen earlier, there are many direct ways in which the DM distribution is connected to marketing literature. Another reason is as follows. Our theoretical framework provides a natural, although simplified, structure to deal with many product categories and to incorporate the entry and exit of brands. We are not aware of any theoretical framework which begins with Gibrat's Law and is able to accomplish the same.

Caveats. We recognize that, like the DM paradigm of brand purchases, our theoretical framework is based on a "reduced-form" stochastic model, and is not directly derived from such "micro" considerations as the perceptions, motivations, and environments of consumers and firms. The same caveats also apply to Gibrat's Law. A topic for future research is the integration of our framework with choice-theoretic models of consumers and firms. Some comments on this topic are presented in the concluding section.

VI. SOME IMPLICATIONS

Managerial implications. One view in marketing is that managers place too much emphasis on market share rather than on other fundamentals of brand performance (Jacobson and Aaker 1985). Our results suggest that the attention managers devote to market share may be justified for high-share brands, where there are large, discrete gaps in the attainable values of market share. This contrasts with the common view, shared both by practitioners and academic researchers, that market shares change in incremental steps in response to small changes in advertising, promotions, pricing, or other activities. If the latter were true, one would expect to find no consistent pattern of the kind we obtain, except that, by definition, a higher-ranked brand will have a larger market share than a lower-ranked brand. One could say nothing about the relationship between the market shares of successively-ranked brands. In contrast, we find a recurrent pattern across product categories thatjumps in market shares are predicted by a power law to a remarkable degree of accuracy. This suggests that each market has a set of "stable" levels of market shares that can be occupied. If this is true, then incremental changes in efforts can produce only short-term changes in market shares. Over time, these will average out and a stable pattern of market shares will emerge. In stable markets, such as those we have examined, these patterns will persist until one firm or another invests large resources to procure the large gains in share that can take it to a higher stable share level. Accordingly, managers seeking sustained gains in market share in stable markets may not be successful by increasing efforts and resources devoted to a brand in small increments.

A related implication concerns the setting of yearly market-share goals, a common task for brand managers. Our results suggest that the smallest target for an increase is to the market share that the power law associates with the next highest rank. Higher possible targets are those associated with higher ranks. These target levels can be useful for planning how much resources to commit to a brand. Note that this approach implies that a manager should primarily consider discrete changes in targets for market share.

A third use of the present results is for assessing the performance of brands, and for relating this assessment to marketing actions. Comparing a brand's current market share against the level predicted for its rank by the power law for the product category can indicate whether it is performing below or above par. If a brand has a higher market share than that predicted by its rank, it might be ready to move up a rank. If it has a lower market share than that predicted by its rank, it might need to be better defended against competitors. This analysis can extend to cross-category comparisons of brands. Thus, a brand with 15% market share in one product category might be a better performer than another brand with a 25% market share in another product category if the first is performing above par for its rank, and the second is not. Managers can use such information when allocating resources across brands in different product categories. Financial analysts can use it to evaluate the performance of brands in different product categories.

A fourth use of the present results is for assessing potential instabilities in markets. For example, if two brands in a product category have the same market shares, then the market is unstable — one or another brand will gain or lose market share until a power-law pattern reappears. This information can be useful for identifying opportunities and potential threats in an unstable market.

Implications for research. The marketing literature on pioneering advantage suggests that earlier entrants sustain larger market shares than later entrants. Kalyanaram, Robinson and Urban (1995) find a power law relation between order of entry and market shares. They estimate the special case of (1) with $a = 0$ and $b = 1/2$, where r is the order of market entry, for prescription anti-ulcer drugs and certain packaged consumer goods. We find that the power law pattern holds for a vastly larger number of product categories without any reference to the order-of-entry. This suggests the need to separate the effects of pioneering advantage from the rank-share relations we report in this paper.

A second implication of our results is that it may be useful to construct models in which marketing efforts and activities are related to the discrete levels of market shares predicted by a power law. The power-law parameters themselves would then be related to such category-level variables as the total advertising expenditures and average prices across brands in the category.

VII. BRIEF REMARKS ON FUTURE RESEARCH

To keep this paper within reasonable bounds of length and scope, and also because of the limitations of our data sets, we have abstracted from several important research issues that are closely related to or that follow naturally from this paper. For example, our data on foods and sporting goods contain market shares aggregated respectively over 120 weeks for a regional market and over one year for a national market, and they do not contain details of individual purchase histories or sales over shorter time spans. In Section III, we described other aspects of our data and also the reasons why our results are likely to hold for comparable aggregations over time spans and geographies. Leaving aside the considerations of these or other data sets, we now present brief remarks on some research issues which are substantive in themselves and which are not limited to testing the boundaries of our findings within proximities of the analysis presented in this paper.

One research issue is whether patterns of the kind that we have identified hold for other aggregations, including those over time (for example, weekly and monthly), geographies (such as states, counties and townships; or rural and urban areas), types of purchase outlets, and consumer segments (such as heavy or light users of a product category, or different benefit segments). Another research issue, distinct from the previous one, is the effect of various aggregations on the values of the estimated parameters. This is because, even if a pattern is known to hold for two or more different data aggregations, the values of the parameters (such as b or g) of this pattern may be similar or dissimilar across these aggregations, depending on the types of aggregations under consideration. Yet another issue, related to the previous two, is whether a pattern under consideration holds over time, and if it does, what is the nature of intertemporal changes in the values of the parameters describing the pattern. Analyses of such research issues are likely to depend in part on the characteristics of the markets and consumer segments which are pertinent to the data under consideration. Among these characteristics are: mature versus new markets (there is some evidence that mature markets exhibit fewer intertemporal changes in the market shares); stable markets versus those in transition; markets for durable versus nondurable goods; seasonality (a large fraction of the annual volume of some products, such as household batteries, is sold within a few weeks of the year); dominance of different types of stores (for some categories, Wal-Mart alone accounts for a large fraction of the total volume); the relative roles of national versus regional or local brands; and the extent to which consumers seek variety. Analyses of the issues of the kind noted in this paragraph will take this paper far afield. For this reason, we view them as topics for future research.

At the beginning of this paper, we noted that our analysis is descriptive rather than causal. The following brief remarks relate this observation to marketing practice and research. At a phenomenological level, firms are concerned, often on an ongoing basis, with variables such as targeting and product positioning, product quality and brand equity, pricing and promotions, advertising expenditures and distribution intensity. A substantial body of research has used such variables to understand the market shares of brands; for example, Guadagni and Little (1983), Lancaster (1990) and McFadden (1986). For brevity of exposition, we will refer to these approaches as "causational." This research also includes, with various degrees of explicitness, considerations such as the histories of firms, the strategic interplay among firms, the behavior of consumers and intermediaries, the dynamics of product growth and innovation, and different kinds of uncertainties and expectations. Much of value has been learned from this literature, and will continue to be learned from its future developments. We believe that our analysis complements the above literature. A macroscopic study such as ours deals with the question of where brands end up in terms of market shares, and not of how they get there. A focus of causational studies is to understand the relationship between the firms' market shares and their efforts and environments. We anticipate these two approaches to converge at some stage in the future.

We conclude with a speculative remark on the role of chance. On the face of it, there is not much in common between the market for soft drinks and the market for handguns in the United States, including the sizes and characteristics of the firms in each of the two markets. We nevertheless find that the power law describes the patterns of market shares in these markets as well as in a number of other disparate product markets considered in this paper. The power law also holds in a number of different ways across product markets, as was shown by our analysis of the derived data sets. This suggests that chance plays a deep role. Kendall (1961, p. 12) has the following to say on the role of chance: "In fact, not only can choice mimic chance, but chance can mimic choice. Consider, for example, a number of persons with equal stakes playing at a fair zero-sum game (A zero-sum game is one in which the losses of some players pass to others, so that no money is lost to the system).[8] Over a period of play, the distribution of their holdings will tend to an unequal pattern of the Pareto type; some people will be reduced to small or zero stakes and a few will accumulate large reserves. And if the game goes on long enough, ultimately all the money will be concentrated in a few hands. Thus chance, which might be expected to level things out, will act in a very inegalitarian way and produce a distribution with a very purposive-looking outcome. It is not Providence but Chance which is on

[8] Evolutions of market shares are, by definition, zero-sum changes; firms' gains and losses must add up to zero. (This footnote is not part of the quotation.)

the side of the big battalions."

APPENDIX

Hierarchical Bayes Estimates of Random-Coefficients Models

In the main body of the paper, we estimated the category-specific parameters for each product category using the data for that category only. These estimates were based on the minimization of the sum of squared errors (MSSE). In this appendix, we present estimates which complement those reported in the paper. We estimate random-coefficients models in which the data are pooled across product categories. Under these models, the parameters for each category are assumed to be draws from a suitable population distribution of the parameters. A caveat regarding the estimates presented here is that the product categories in our data are not a random sample of the respective universes of product categories of foods or sporting goods.

For convenience, we recall our basic notation. We use the index $i = 1$ to m to represent product categories. Within the product category i, there are n_i brands. We assign the index $j = 1$ to n_i to the brands in category i in non-increasing order of market shares. We use the protocol described in the paper to assign a rank r_{ij} to brand j in category i. The market share of brand j in product category i is s_{ij}. The expressions for the power law and the exponential form are then respectively

$$s_{ij} = A_i(a_i + r_{ij})^{-b_i}; \text{and} \tag{A1}$$

$$s_{ij} = G_i e^{-g_i r_{ij}}. \tag{A2}$$

Our objective is to estimate the parameters (a_i, b_i, g_i); for reasons described in the text, we do not estimate the parameters (A_i, G_i). For later use, we define $z_{ij} \equiv s_{i1}/s_{ij}$. Then, from ($A1$) and ($A2$)

$$z_{ij} = \left(\frac{a_i + r_{ij}}{a_i + r_{i1}}\right)^{b_i} \tag{A3}$$

for the power law, and

$$z_{ij} = e^{-g_i(r_{i1} - r_{ij})} \tag{A4}$$

for the exponential form. In each of our product categories, only one brand has the highest market share. Hence $r_{i1} = 1$, $z_{i1} = 1$ and $r_{ij} \geq 2$ for all $2 \leq j \leq n_i$ and all $1 \leq i \leq m$. Since all z_{i1} equal unity, we do not use them in our estimates. Define $x_{ij} \equiv r_{ij} - 1$, and $y_{ij} \equiv \ln z_{ij}$. We take logarithms of both sides of ($A3$) and ($A4$) to obtain the following expressions for the power law and the exponential form respectively:

$$y_{ij} = b_i \ln\left(1 + \frac{x_{ij}}{a_i + 1}\right), \quad \text{and} \tag{A5}$$

$$y_{ij} = g_i x_{ij}. \tag{A6}$$

We use these expressions to estimate the parameters (a_i, b_i, g_i). We now describe the key aspects of our method of estimation, first for the power law and then for the exponential form.

Power law. We assume that the values of a_i and b_i for each product category are random draws from a suitable population distribution. In many situations, it is natural to assume a normal distribution for such parameters. However, for reasons noted in the paper, we require that $a_i + 1 > 0$ and $b_i > 0$. Therefore, a reasonable assumption, which we adopt here, is that $a_i + 1$ and b_i have a bivariate lognormal distribution. That is, $\beta_{i1} \equiv \ln(a_i + 1)$ and $\beta_{i2} \equiv \ln(b_i)$ have a bivariate normal distribution. Let $\boldsymbol{\beta}_i \equiv (\beta_{i1}, \beta_{i2})$ be drawn from a population with distribution $N(\boldsymbol{\mu}, \boldsymbol{\Lambda})$, where $\boldsymbol{\mu}$ is a 2×1 vector of population means and $\boldsymbol{\Lambda}$ is a 2×2 covariance matrix. We estimate $\boldsymbol{\beta_i}$ in the model

$$y_{ij} = \exp(\beta_{i2}) \ln \left[1 + \frac{x_{ij}}{\exp(\beta_{i1})} \right] + e_{ij} \quad \text{for } 2 \leq j \leq n_i \text{ and } 1 \leq i \leq m, \qquad (A7)$$

where $e_{ij} \sim N(0, \sigma^2)$. We define the vectors $\mathbf{y_i} \equiv (y_{i1}, y_{i2}, y_{i3} \ldots, y_{in_i})$ and $\mathbf{y} \equiv (\mathbf{y_1}, \mathbf{y_2}, \mathbf{y_3}, \ldots, \mathbf{y_m})$. The likelihood function for product category i is

$$L_i(\boldsymbol{\beta_i}) = \prod_{j=2}^{n_i} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(y_{ij} - w_{ij})^2}{\sigma^2} \right) \quad \text{for } 1 \leq i \leq m, \qquad (A8)$$

where $w_{ij} \equiv \exp(\beta_{i2}) \ln(1 + (x_{ij}/\exp(\beta_{i1})))$. Let $\phi(\boldsymbol{\beta}_i)$ denote the density for the bivariate normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Lambda})$. We assume that the data for each product category is independent of the data for the other product categories. The unconditional likelihood for a random sample of m categories is given by the continuous mixture

$$L = \prod_{i=1}^{m} \int \int L_i\left(\boldsymbol{\beta}_i\right) \phi(\boldsymbol{\beta}_i) d\boldsymbol{\beta}_i. \qquad (A9)$$

The unconditional likelihood cannot be written in closed form because the normal population distribution is not conjugate to the conditional likelihood L_i. We therefore use hierarchical Bayes methods to estimate the parameters. We use proper but diffuse priors. The joint prior is a product of independent priors over $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$ and σ^2. We assume that the prior for $\boldsymbol{\mu}$ is a normal distribution $N(\boldsymbol{\delta}, \boldsymbol{C})$. The covariance matrix $\boldsymbol{C}$ may be specified to be diagonal with the elements (variances) set to a large value (we set this value to 100) to represent vague knowledge. Under this assumption for $\boldsymbol{C}$, the value of $\boldsymbol{\delta}$ is no longer critical and so we set $\boldsymbol{\delta} = 0$. We assume that the prior for the precision matrix $\boldsymbol{\Lambda}^{-1}$ has a Wishart distribution[9] $W(\rho, (\rho\boldsymbol{R})^{-1})$ with

[9]The Wishart density with ν degrees of freedom is

$$p(\boldsymbol{V}|\nu, \boldsymbol{\Omega}) = c\frac{|\boldsymbol{V}|^{(\nu-k-1)/2}}{|\boldsymbol{\Omega}|^{\nu/2}} \exp\left(-\operatorname{tr}(\boldsymbol{\Omega}^{-1}\boldsymbol{V})/2 \right),$$

$\rho \geq 2$ degrees of freedom, where $\boldsymbol{R}$ is a 2×2 symmetric positive definite matrix. The parameterization $\boldsymbol{\Lambda}^{-1} \sim W(\rho, (\rho\boldsymbol{R})^{-1})$ implies that $E(\boldsymbol{\Lambda}^{-1}) = \boldsymbol{R}^{-1}$. Hence, $\boldsymbol{R}$ is approximately the expected prior covariance matrix of the individual-specific $\boldsymbol{\beta}_i$'s. As $\text{Var}(\Lambda_{ij})$ is decreasing in ρ, small values of ρ correspond to vaguer prior distributions. We set $\boldsymbol{R}$ to be an identity matrix and $\rho = 3$. Finally, we assume that the prior for σ^2 is Inverse Gamma $IG(a, b)$, with $a = 3$ and $b = 1$.

We use Markov Chain Monte Carlo (MCMC) methods to simulate dependent draws from the joint posterior distribution. This approach replaces one complicated draw from the posterior distribution with a series of relatively simple draws from distributions that are easy to sample. Samples from the posterior are obtained by iteratively sampling from the full conditional distributions of different parameter blocks. The MCMC sampler is run for a large number of iterations. This iterative scheme generates a Markov chain that converges in distribution to the joint posterior distribution under fairly general conditions (Tierney 1994). After an initial transient phase, also known as the burn-in period, the chain converges to the posterior distribution of parameters; all subsequent draws may be regarded as sample draws from the posterior distribution. After this convergence takes place, the sample of draws can be used to approximate the posterior to any desired degree of accuracy.

If all full conditionals have closed forms, the MCMC sampler reduces to the Gibbs sampling procedure (Gelfand and Smith 1990; Geman and Geman 1984). In many situations though, the full conditional distributions for certain parameters are known only up to a normalizing constant. For these parameters, the Metropolis–Hastings step (Metropolis et al. 1953, Hastings 1970, Chib and Greenberg 1995) can be used.

In the present context, we need to generate random draws for the unknowns $\{\{\boldsymbol{\beta}_i\}, \boldsymbol{\mu}, \boldsymbol{\Lambda}, \sigma^2\}$. Each iteration of the MCMC sampler involves sampling from the full conditionals associated with each block of parameters. The sampler produces draws for the category-specific parameters $\{\boldsymbol{\beta}_i\}$, and therefore allows for a proper accounting of the uncertainty regarding these parameters. The $(S+1)$-st iteration of the MCMC method involves generating random draws using the following full conditional distributions:

(a) The full conditional for the category-specific parameters $\boldsymbol{\beta}_i$ cannot be written in closed form because the population distribution $N(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ is not conjugate to the category-level power-law likelihood. We therefore use a Metropolis–Hastings

where $\boldsymbol{V}$ and $\boldsymbol{\Omega}$ are square, symmetric and positive definite matrices, each with $k \leq \nu$ rows and columns. The symbol tr denotes the trace of a matrix. The proportionality constant c has the value

$$c = \frac{1}{2^{(\nu k)/2} \pi^{k(k-1)/4} \prod_{j=1}^{k} \Gamma\left(\frac{\nu+1-j}{2}\right)},$$

where $\Gamma(\cdot)$ denotes the Gamma function. Let Ω_{ij} denote the ij-th element of $\boldsymbol{\Omega}$, and let V_{ij} denote the ij-th element of $\boldsymbol{V}$. The elements of $\boldsymbol{\Omega}$ are scale parameters. The means, variances and covariances of the elements of $\boldsymbol{V}$ are given by $E(V_{ij}) = \nu\Omega_{ij}$, $\text{Var}(V_{ij}) = \nu(\Omega_{ij}^2 + \Omega_{ii}\Omega_{jj})$, and $\text{Cov}(V_{ij}, V_{kl}) = \nu(\Omega_{ik}\Omega_{jl} + \Omega_{il}\Omega_{jk})$, respectively.

step to generate draws from this full conditional. For category i, the posterior density is proportional to the likelihood

$$\prod_{j=2}^{n_i} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(y_{ij}-w_{ij})^2}{\sigma^2}\right) \tag{A11}$$

and the prior $N(\boldsymbol{\mu}, \boldsymbol{\Lambda})$. We use a random walk Metropolis step to generate draws for $\boldsymbol{\beta}_i$. This requires generating a candidate $\boldsymbol{\beta}_i^C$ from a multivariate normal proposal density $N(\boldsymbol{\beta}_i^S, \boldsymbol{\Omega}_i)$. The proposal density is centered on the old value of $\boldsymbol{\beta}_i^S$, from iteration S. The variance of the proposal density, $\boldsymbol{\Omega}_i$, also known as the tuning constant, is set to allow rapid mixing of the chain. The generated candidate $\boldsymbol{\beta}_i^C$ is accepted with the acceptance probability

$$\alpha(\boldsymbol{\beta}_i^S, \boldsymbol{\beta}_i^C) = \min\left\{1, \frac{L(\boldsymbol{\beta}_i^C)\phi(\boldsymbol{\beta}_i^C|\boldsymbol{\mu}, \boldsymbol{\Lambda})}{L(\boldsymbol{\beta}_i^S)\phi(\boldsymbol{\beta}_i^S|\boldsymbol{\mu}, \boldsymbol{\Lambda})}\right\}, \tag{A12}$$

where $\phi(\cdot)$ represent the normal density. If the candidate is accepted, $\boldsymbol{\beta}_i^{S+1} = \boldsymbol{\beta}_i^C$; otherwise, $\boldsymbol{\beta}_i^{S+1} = \boldsymbol{\beta}_i^S$. The parameters for the different product categories can be drawn in sequence. As the acceptance probability depends only on the ratio of the posterior densities, any normalizing constant cancels out. Hence the Metropolis step can be used in instances where the full conditional is not completely known.

(b) The full conditional for $\boldsymbol{\mu}$ is a normal distribution. The prior $N(\boldsymbol{\delta}, \boldsymbol{C})$ is conjugate to the population distribution, $\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Lambda})$. The posterior full conditional distribution can be written as

$$p(\boldsymbol{\mu}|\{\boldsymbol{\beta}_i\}, \boldsymbol{\Lambda}) = N(\tilde{\boldsymbol{\mu}}, \boldsymbol{V}_\mu), \tag{A13}$$

with posterior precision $\boldsymbol{V}_\mu^{-1} = \boldsymbol{C}^{-1} + m\boldsymbol{\Lambda}^{-1}$ and posterior mean $\tilde{\boldsymbol{\mu}} = \boldsymbol{V}_\mu(\boldsymbol{C}^{-1}\boldsymbol{\delta} + \sum_{i=1}^m \boldsymbol{\Lambda}^{-1}\boldsymbol{\beta}_i)$.

(c) As the prior $W(\rho, (\rho\boldsymbol{R})^{-1})$ is conjugate to the normal population distribution of the category-specific coefficients, the full conditional distribution for the population precision matrix $\boldsymbol{\Lambda}^{-1}$ is Wishart. The full conditional can be written as

$$p(\boldsymbol{\Lambda}^{-1}|\{\boldsymbol{\beta}_i\}, \boldsymbol{\mu}) = W\Big(\rho + m, \Big[\sum_{i=1}^m (\boldsymbol{\beta}_i - \boldsymbol{\mu})(\boldsymbol{\beta}_i - \boldsymbol{\mu})' + \rho\boldsymbol{R}\Big]^{-1}\Big). \tag{A14}$$

(d) The full conditional for the error variance is an Inverse Gamma distribution given by

$$p(\sigma^2|\boldsymbol{y}, \boldsymbol{\beta}_i) = IG\left[a + \frac{N}{2}, \left(\sum_{i=1}^m \sum_{j=2}^{n_i} \frac{(y_{ij}-w_{ij})^2}{2} + b^{-1}\right)^{-1}\right], \tag{A15}$$

where N is the number of observations.

Exponential form. We omit some details of our estimation methods for the exponential form because these details are identical to those, described above, for the power law. We assume that the values of g_i for each category are random draws from a suitable probability distribution. For reasons stated in the paper, we require $g_i > 0$ in ($A6$). Accordingly, we assume that g_i has a lognormal distribution. That is, $\beta_i \equiv \ln g_i$ has a normal distribution with population mean μ and variance τ^2. We estimate β_i in the model

$$y_{ij} = \exp(\beta_i) \cdot x_{ij} + e_{ij}, \quad \text{for } 2 \leq j \leq n_i \text{ and } 1 \leq i \leq m, \qquad (A16)$$

where $e_{ij} \sim N(0, \sigma^2)$. The likelihood function for product category i is

$$L_i(\beta_i) = \prod_{j=2}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(y_{ij} - w_{ij})^2}{\sigma^2}\right), \text{ for } 1 \leq i \leq m, \qquad (A17)$$

where $w_{ij} \equiv \exp(\beta_i) \cdot x_{ij}$. Let $\phi(\beta_i)$ denote the density for the normal distribution $N(\mu, \tau^2)$. The unconditional likelihood for a random sample of m categories is given by the continuous mixture

$$L = \prod_{i=1}^{m} \int L_i(\beta_i)\, \phi(\beta_i) d\beta_i. \qquad (A18)$$

The normal population distribution is not conjugate to the conditional likelihood L_i. Therefore the unconditional likelihood cannot be written in closed form. As with the power law, we use hierarchical Bayes methods to estimate the parameters. We use (proper) independent and diffuse priors over μ, τ^2 and σ^2. We assume that the prior for μ is a normal distribution $N(\delta, c)$. We set $\delta = 0$ and $c = 100$ to represent vague knowledge. We assume that the prior distribution of τ^2 is Inverse Gamma $IG(a_1, b_1)$, and that the prior distribution of σ^2 is Inverse Gamma $IG(a_2, b_2)$. We set $a_1 = a_2 = 3$ and $b_1 = b_2 = 1$.

As in the case of the power law, we use MCMC methods to simulate dependent draws from the joint posterior distribution. The $(S + 1)$-st iteration of the MCMC method involves generating random draws using the following full conditional distributions:

(a) The full conditional for τ^2 is an Inverse Gamma distribution given by

$$p(\tau^2 | \{\beta_i\}, \mu) = IG\left[a_1 + \frac{m}{2}, \left(\sum_{i=1}^{m} \frac{(\beta_i - \mu)^2}{2} + \frac{1}{b_1}\right)^{-1}\right], \qquad (A19)$$

where m is the number of product categories.

(b) The full conditional for σ^2 is an Inverse Gamma distribution given by

$$p(\sigma^2) = IG\left[a_2 + \frac{N}{2}, \left(\sum_{i=1}^{m}\sum_{j=2}^{n_i}\frac{(y_{ij}-w_{ij})^2}{2} + \frac{1}{b_2}\right)^{-1}\right]. \qquad (A20)$$

(c) The full conditional for μ is a normal distribution. As the prior $N(\delta, c)$ is conjugate to the population distribution $\beta_i \sim N(\mu, \tau^2)$, the posterior full conditional distribution can be written as

$$p(\mu|\{\beta_i\}, \tau) = N(\tilde{\mu}, v_\mu), \qquad (A21)$$

with posterior precision $v_\mu^{-1} = c^{-1} + m\tau^{-2}$ and posterior mean $\tilde{\mu} = v_\mu(c^{-1}\delta + m\bar{\beta}\tau^{-2})$, where $\bar{\beta} \equiv \sum_{i=1}^{m}\beta_i/m$.

(d) The full conditional for β_i is a normal distribution given by

$$p(\beta_i|\mu, \tau^2, \{y_i\}, \sigma^2) = N(\hat{\beta}_i, v_{\beta i}) \qquad (A22)$$

where $v_{\beta_i}^{-1} = \tau^{-2} + \sum_{j=2}^{n_i} x_{ij}^2\sigma^{-2}$ and $\hat{\beta}_i = v_{\beta i}[\tau^{-2}\mu + \sum_{j=2}^{n_i} x_{ij}\sigma^{-2}y_{ij}]$.

Results and comparisons of models. For the power law, $a = \exp(\beta_1) - 1$ and $b = \exp(\beta_2)$, where β_1 and β_2 are normally distributed random variables. Following Press (1972, p. 139, Theorem 6.4.1)

$$E[\exp(\beta_j)] = \exp[E(\beta_j) + \frac{1}{2}\mathrm{var}(\beta_j)], \; j = 1, 2, \qquad (A23)$$

where $\mathrm{var}(\beta_j)$ is the variance of β_j. For the exponential form, $g = \exp(\beta)$, where β has a normal distribution. Hence $E[g] = E[\exp(\beta_j)]$ is given by an expression analogous to (*A*23). For the food categories, the estimated means and variances of β_1, β_2 and β are

$$E[\beta_1] = 1.72, E[\beta_2] = 1.43, E[\beta] = -0.92, \hat{\sigma}_1^2 = 2.53, \hat{\sigma}_2^2 = 1.39, \text{ and } \hat{\sigma}^2 = 0.36.$$

From these we obtain

$$E[a] = 19.79, E[b] = 8.37, \text{ and } E[g] = 0.48.$$

Correspondingly, for the sporting-goods categories,

$$E[\beta_1] = 0.84, E[\beta_2] = 0.62, E[\beta_3] = -1.22, \hat{\sigma}_1^2 = 1.92, \hat{\sigma}_2^2 = 0.79, \text{ and } \hat{\sigma}^2 = 0.38.$$

From these we obtain

$$E[a] = 6.05, E[b] = 2.76, \text{ and } E[g] = 0.36.$$

As in the MSSE estimates presented in the main body of the paper, there are more food categories with larger estimates of a and b, and this is reflected in the above sets of mean values.

Let M_p denote the model for the power law and let M_e denote the model for the exponential form. We use Bayes factors (Kass and Raftery 1995) to compare the fits of M_p and M_e. Let $\boldsymbol{\eta}_l$ denote the parameter vector, let $\pi_l(\boldsymbol{\eta}_l|M_l)$ denote the prior, and let $f(\boldsymbol{y}|M_l, \boldsymbol{\eta}_l)$ denote the sampling density for model $l \in \{p, e\}$. The posterior odds ratio can be written as

$$\frac{\Pr(M_p|\boldsymbol{y})}{\Pr(M_e|\boldsymbol{y})} = \frac{\Pr(M_p)}{\Pr(M_e)} \times \frac{m(\boldsymbol{y}|M_p)}{m(\boldsymbol{y}|M_e)}, \tag{A24}$$

where $m(\boldsymbol{y}|M_l) = \int f(\boldsymbol{y}|M_l, \boldsymbol{\eta}_l)\pi_l(\boldsymbol{\eta}_l|M_l)d\boldsymbol{\eta}_l$ is the marginal likelihood of model l. The Bayes factor is the ratio of values of the marginal likelihoods. These values can be computed for each model based on the MCMC draws, using an importance sampling scheme outlined in Newton and Raftery (1994). The values of the log marginal likelihoods are as follows:

Data	Model	Log marginal likelihood
Food categories	Power law	–35.84
Food categories	Exponential form	–299.50
Sporting-goods categories	Power law	–187.14
Sporting-goods categories	Exponential form	–610.18

The Bayes factor has a value of e^{263} for the food categories and e^{423} for the sporting goods categories. This suggests an overwhelmingly better fit for the power law than for the exponential form.

REFERENCES

Adamic, L.A. and B.A. Huberman (2000), "The Nature of Markets in the World Wide Web," *Quarterly Journal of Electronic Commerce,* 1, 5–12.

Aoki, M. (1996), *New Approaches to Macroeconomic Modeling,* New York, NY: Cambridge University Press.

Bass, F.M. (1995), "Empirical Generalizations and Marketing: A Personal View," *Marketing Science,* 14 (3) Part 2, G6–G19.

Bass, F.M. (1969), "A New Product Growth Model for Consumer Durables," *Management Science,* 15 (1), 215–227.

Boston Consulting Group (1987), "The Rule of Three and Four," *Perspectives,* 187, Boston, MA.

Brock, W.A. (1999), "Scaling in Economics: A Reader's Guide," *Industrial and Corporate Change,* 8 (3), 409–446.

Buzzell, R.D. (1981), "Are There 'Natural' Market Structures?" *Journal of Marketing,* 45 (1), 42–52.

Chen, W.-C. (1978), "On Zipf's Law," Ph.D. dissertation, University of Michigan, Ann Arbor, MI.

Chib, S. and E. Greenberg (1995), "Understanding the Metropolis–Hastings Algorithm," *The American Statistician*, 49, 327–335.

Christensen, L.R., D.W. Jorgenson and L.J. Lau (1975), "Transcendental Logarithmic Utility Functions," *American Economic Review,* 65 (3), 367–383.

Chung, K.H. and R.A.K. Cox (1994), "A Stochastic Model of Superstardom: An Application of the Yule Distribution," *Review of Economics and Statistics,* 76 (4), 771–775.

Deaton, A. and J. Muellbauer (1980), *Economics and Consumer Behavior,* Cambridge, UK: Cambridge University Press.

Ehrenberg, A.S.C. (1995), "Empirical Generalizations, Theory and Methods," *Marketing Science,* 14 (3) Part 2, G20–G28.

Ehrenberg, A.S.C. (1982), "Lawlike Relationships," *Research in Marketing Series,* No. 82/1, London Business School.

Fader, P. and D.C. Schmittlein (1993), "Excessive Behavioral Loyalty for High-Share Brands: Deviations From the Dirichlet Model for Repeat Buying," *Journal of Marketing Research,* 30 (4), 478–493.

Feller, W. (1968), *An Introduction to Probability Theory and Its Application*, Vol. 1, 3rd ed., New York, NY: Wiley.

Gabaix, X. (1999), "Zipf's Law for Cities," *Quarterly Journal of Economics,* 114 (3), 738–767.

Gell-Mann, M. (1994), *The Quark and the Jaguar: Adventures in the Simple and the Complex*, New York: W. H. Freeman.

Gelfand, A.E. and A.F.M. Smith (1990), "Sampling-based approaches to calculating marginal densities," *Journal of the American Statistical Association*, 85, 398–409.

Geman, S. and D. Geman (1984), "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721–741.

Goodhardt, G.J., A.S.C. Ehrenberg and C. Chatfield (1984), "The Dirichlet: A Comprehensive Model of Buying Behavior," *Journal of the Royal Statistical Society A,* 147 (5), 621–655.

Guadagni, P.M. and J.D.C. Little (1983), "Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science,* 2 (3), 203–238.

Hastings, W.K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and their Applications," *Biometrika*, 57, 97–109.

Hill, B.M. (1974), "The Rank-Frequency Form of Zipf's Law," *Journal of the American Statistical Association,* 69 (348), 1017–1026.

Hill, B.M. (1970) "Stronger Forms of Zipf's Law and Prior Distributions for the Composition of a Population," *Journal of the American Statistical Association,* 65 (331), 1220–1232.

Hill, B.M. and M. Woodroofe (1975), "Stronger Forms of Zipf's Law," *Journal of the American Statistical Association,* 70 (349), 212–219.

Ijiri, Y. and H.A. Simon (1975), "Some Distributions Associated with Bose–Einstein Statistics," *Proceedings of the National Academy of Sciences,* 72 (5), 1654–1657.

Kalyanaram, G., W.T. Robinson and G.L. Urban (1995), "Order of Market Entry: Established Empirical Generalizations, Emerging Empirical Generalizations, and Future Research," *Marketing Science,* 14 (3) Part 2, G212–G221.

Kass, R.E. and A.E. Raftery (1995), "Bayes Factor," *Journal of the American Statistical Association*, 90, 773–795.

Kendall, M. G. (1961), "Presidential Address," *Journal of the Royal Statistical Society: Series A (General),* 124 (1), 1–16.

Kotler, P. (1977), *Marketing Management: Analysis, Planning, Implementation and Control,* Englewood Cliffs, NJ: Prentice Hall.

Lancaster, K. (1990), "The Economics of Product Variety: A Survey," *Marketing Science,* 3 (9), 189–206.

Mandelbrot, B. (1963), "New Methods in Statistical Economics," *Journal of Political Economy,* 71 (5), 421–440.

McFadden, D. (1986), "The Choice Theory Approach to Market Research," *Marketing Science,* 5 (4), 275–297.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller (1953), "Equations of State Calculations by Fast Computing Machines," *Journal of Chemical Physics*, 21, 1087–1092.

Newton, M.A. and A.E. Raftery (1994), "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society, Series B,* 3, 3–48.

Persky, J. (1992), "Pareto's Law," *Journal of Economic Perspectives,* 6 (2), 181–192.

Press, S.J. (1972), *Applied Multivariate Statistics,* New York, NY: Holt, Rinehart and Winston.

Simon, H.A. (1968), "On Judging the Plausibility of Theories," in B. van Rootselaar and J. F. Staal (Eds.), *Logic, Methodology and Philosophy of Sciences III,* Amsterdam: North Holland, 439–459.

Tierney, L. (1994), "Markov Chains for Exploring Posterior Distributions (with Discussions)," *Annals of Statistics*, 22, 1701–1762.

TABLE 1

Estimates for each product category: Foods

No. of	*Exponential Form*		*Power Law*		
brands	*g*	R^2	*a*	*b*	R^2
27	**0.06**	0.84	-0.17	**0.59**	0.97
22	**0.11**	0.91	1.01	**1.10**	0.97
21	**0.12**	0.84	**-0.65**	**0.75**	0.91
21	**0.12**	0.97	16.92	**3.35**	0.98
20	**0.15**	0.91	**1.22**	**1.47**	0.99
20	**0.15**	0.95	**4.78**	**2.03**	0.99
20	**0.12**	0.92	2.35	**1.31**	0.95
19	**0.13**	0.89	0.40	**1.06**	0.99
17	**0.18**	0.96	**3.66**	**2.00**	0.99
16	**0.18**	0.98	17.04	**4.52**	0.99
16	**0.20**	0.97	9.56	**3.41**	0.98
13	**0.25**	0.90	0.87	**1.63**	0.98
13	**0.23**	0.96	4.40	**2.38**	0.98
12	**0.20**	0.90	0.75	**1.23**	0.98
12	**0.24**	0.90	0.31	**1.33**	0.98
12	**0.29**	0.99	**28.58**	10.00	0.99
11	**0.30**	0.94	**28.08**	10.00	0.94
11	**0.30**	0.99	**27.33**	10.00	0.99
11	**0.36**	0.94	1.98	**2.57**	0.98
11	**0.28**	0.90	0.89	**1.62**	0.94
10	**0.38**	0.83	**-0.61**	**1.28**	0.97
10	**0.40**	0.96	**19.85**	10.00	0.96
10	**0.41**	0.97	**19.29**	10.00	0.97
9	**0.35**	0.93	4.57	3.20	0.93

No. of	*Exponential Form*		*Power Law*		
brands	*g*	R^2	*a*	*b*	R^2
8	**0.55**	0.92	2.92	3.82	0.95
8	**0.50**	0.97	8.29	6.31	0.98
8	**0.48**	0.95	**16.85**	10.00	0.95
8	**0.49**	0.98	15.00	9.39	0.98
7	**0.32**	0.92	**27.89**	10.00	0.92
7	**0.51**	0.98	**16.07**	10.00	0.98
7	**0.61**	0.85	3.79	4.60	0.87
7	**0.68**	0.96	**10.97**	10.00	0.96
6	**0.55**	0.96	**15.00**	10.00	0.96
6	**0.86**	0.95	3.41	5.72	0.97
6	**0.78**	0.96	1.93	4.00	0.99
6	**0.55**	0.99	12.36	8.67	0.99
6	**0.71**	0.91	-0.16	**1.94**	0.99
6	**0.81**	0.86	**9.45**	10.00	0.86
5	**0.92**	0.94	0.22	2.57	0.97
5	**0.85**	1.00	**8.86**	10.00	1.00
5	0.99	0.69	**-0.99**	**0.75**	1.00
5	**0.78**	0.89	0.02	1.99	0.91
5	**0.74**	0.92	**10.71**	10.00	0.92
5	**1.01**	0.97	**7.26**	10.00	0.97
4	0.59	0.69	15.02	10.00	0.68
4	0.55	0.83	-0.91	0.48	0.96
4	**1.19**	0.99	**6.12**	10.00	0.99
4	0.77	0.83	10.94	10.00	0.82

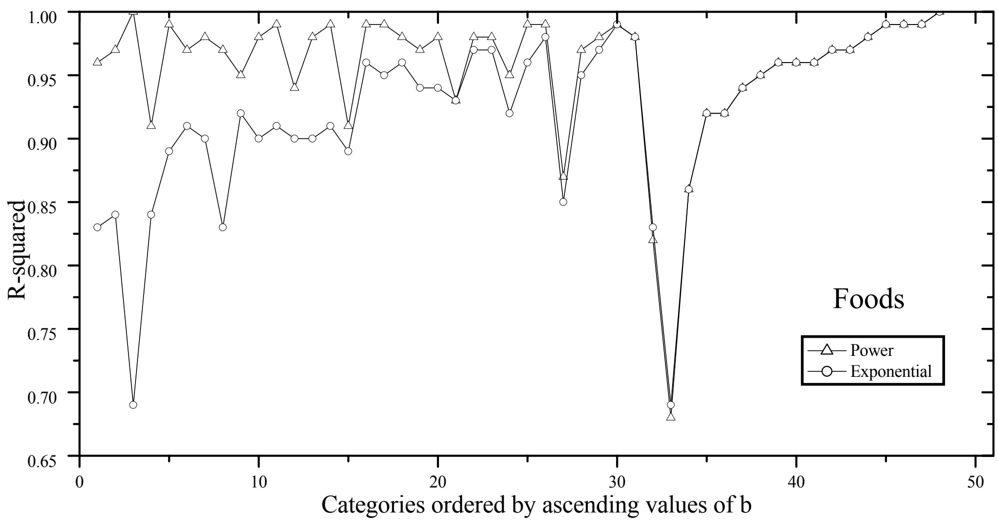
TABLE 2

Estimates for each product category: Sporting goods

Product Category	*No. of brands*	*Exponential Form*		*Power Law*		
		g	R^2	*a*	*b*	R^2
Sports sandals	37	**0.07**	0.92	**3.37**	**1.29**	0.99
Boat shoes	37	**0.07**	0.94	**3.95**	**1.32**	0.99
Backpacks	36	**0.09**	0.92	**3.29**	**1.62**	0.96
Walking shoes	34	**0.09**	0.86	**0.61**	**1.22**	0.99
Hiking boots	34	**0.08**	0.89	**1.42**	**1.10**	0.98
Sneakers (gym shoes)	31	**0.10**	0.85	**0.48**	**1.21**	0.99
Fitness shoes	30	**0.12**	0.88	0.38	**1.37**	0.99
Tennis shoes	27	**0.12**	0.86	0.28	**1.24**	0.99
Scooters	26	**0.12**	0.75	**-0.88**	**0.85**	0.97
Aerobic shoes	24	**0.15**	0.83	-0.02	**1.30**	0.96
Golf clubs	19	**0.14**	0.94	**2.24**	**1.49**	1.00
Soccer balls	18	**0.17**	0.95	**4.78**	**2.25**	0.98
Cross-training shoes	18	**0.23**	0.86	0.28	**1.76**	0.99
Sleeping bags	16	**0.18**	0.72	**-0.97**	**0.63**	0.95
Golf club sets	16	**0.14**	0.98	21.73	4.06	0.98
Jogging shoes	16	**0.28**	0.91	2.67	**2.81**	0.97
Golf bags	15	**0.17**	0.99	**12.53**	**3.25**	0.99
Skateboarding shoes	15	**0.22**	0.79	**-0.56**	**1.09**	0.98
Hunting boots	15	**0.13**	0.87	0.08	**0.77**	0.97
Soccer shoes	13	**0.32**	0.88	0.05	**1.73**	0.98
Fishing reels	13	**0.24**	0.97	16.67	5.60	0.98
Basketball shoes	12	**0.42**	0.89	0.43	**2.33**	0.97

Product Category	*No. of brands*	*Exponential Form*		*Power Law*		
		g	R^2	*a*	*b*	R^2
Baseball shoes	12	**0.35**	0.87	0.08	**1.79**	0.97
Tents	11	**0.31**	0.93	0.46	**1.63**	0.98
Rifles	9	**0.31**	0.99	**27.67**	10.00	0.98
Water-sport shoes	9	**0.14**	0.95	**64.53**	10.00	0.95
Golf shoes	9	**0.42**	0.94	0.77	**2.08**	0.99
Treadmills	8	**0.43**	0.84	**-0.68**	**1.07**	0.99
Ice hockey sticks	8	**0.27**	0.98	**32.64**	10.00	0.98
Handguns	8	**0.30**	0.97	**29.03**	10.00	0.97
Cycling shoes	8	**0.23**	1.00	**39.39**	10.00	1.00
Bowling shoes	8	**0.54**	0.90	0.46	2.03	0.94
Billiard pool sticks	8	**0.19**	0.95	-0.07	**0.59**	1.00
Baseball/Softball bats	8	**0.37**	0.94	**22.69**	10.00	0.93
Baseball mitts	8	**0.50**	0.95	**15.58**	10.00	0.95
Tennis racquets	7	**0.53**	0.91	-0.02	**1.73**	0.98
Track shoes	7	**0.50**	0.95	0.09	**1.63**	1.00
Cheerleading shoes	7	**0.28**	0.66	**-0.98**	**0.37**	0.99
Basketballs	7	**0.65**	0.95	11.20	9.80	0.96
Volleyballs	6	**0.25**	0.97	0.48	0.84	1.00
Footballs	6	**0.62**	0.91	0.11	1.89	0.97
Football shoes	5	**0.95**	0.95	1.97	4.42	0.96
Air pistols	4	**1.29**	0.91	**32.64**	10.00	0.98

FIGURE 1: R-squared values for each product category



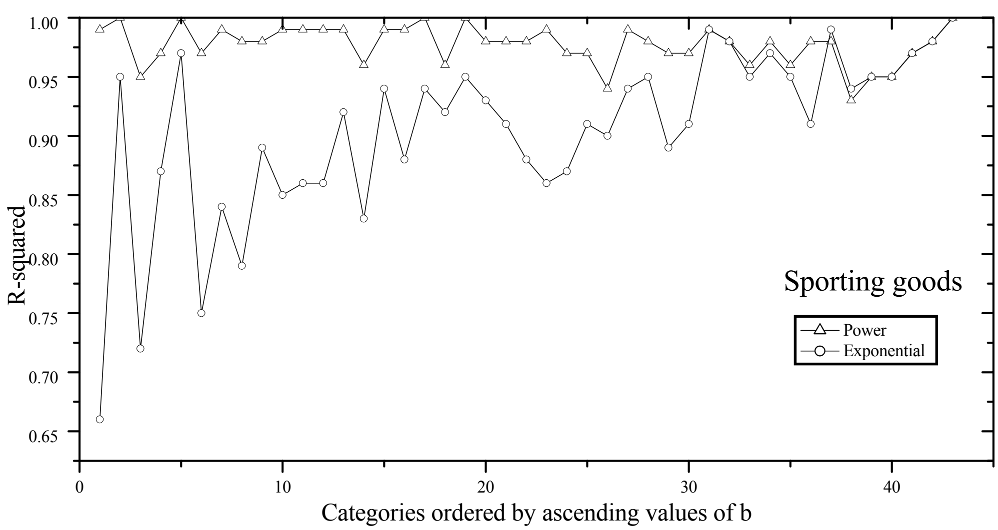


FIGURE 2: Market shares by rank, across product categories

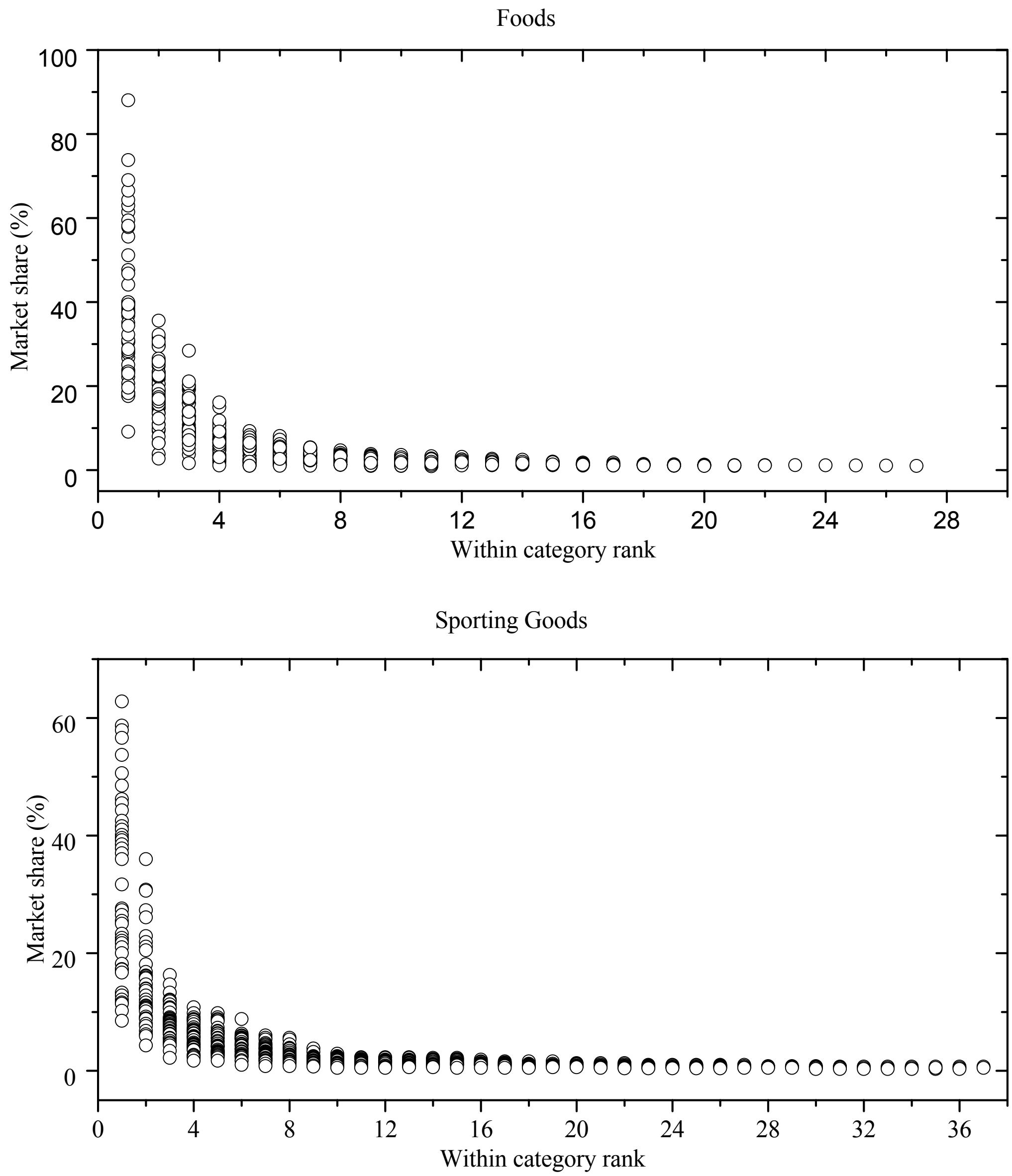


FIGURE 3: Market share versus rank across 506 food brands and 665 sporting goods brands

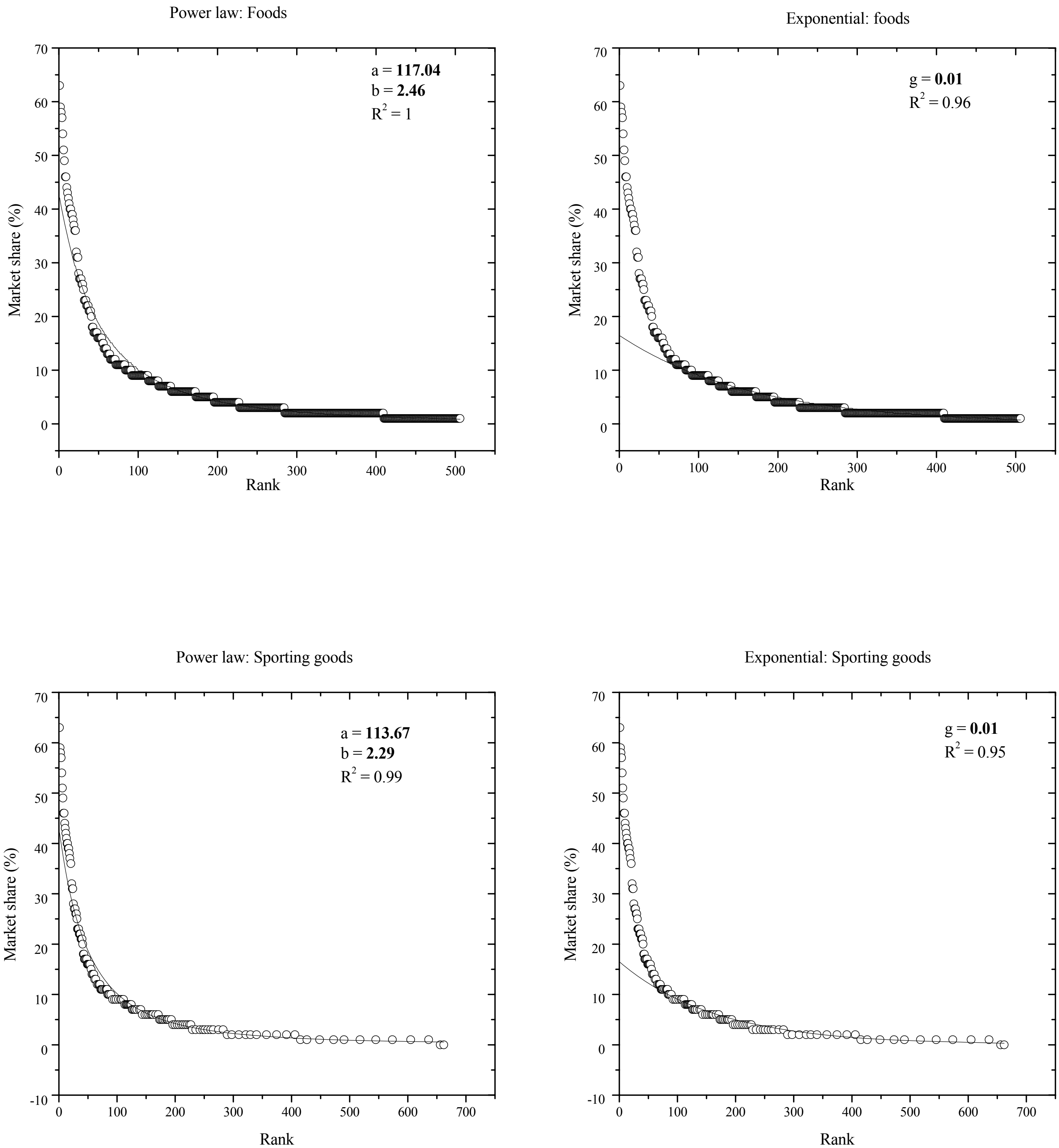


FIGURE 4: Highest-ranked market shares, across product categories of foods

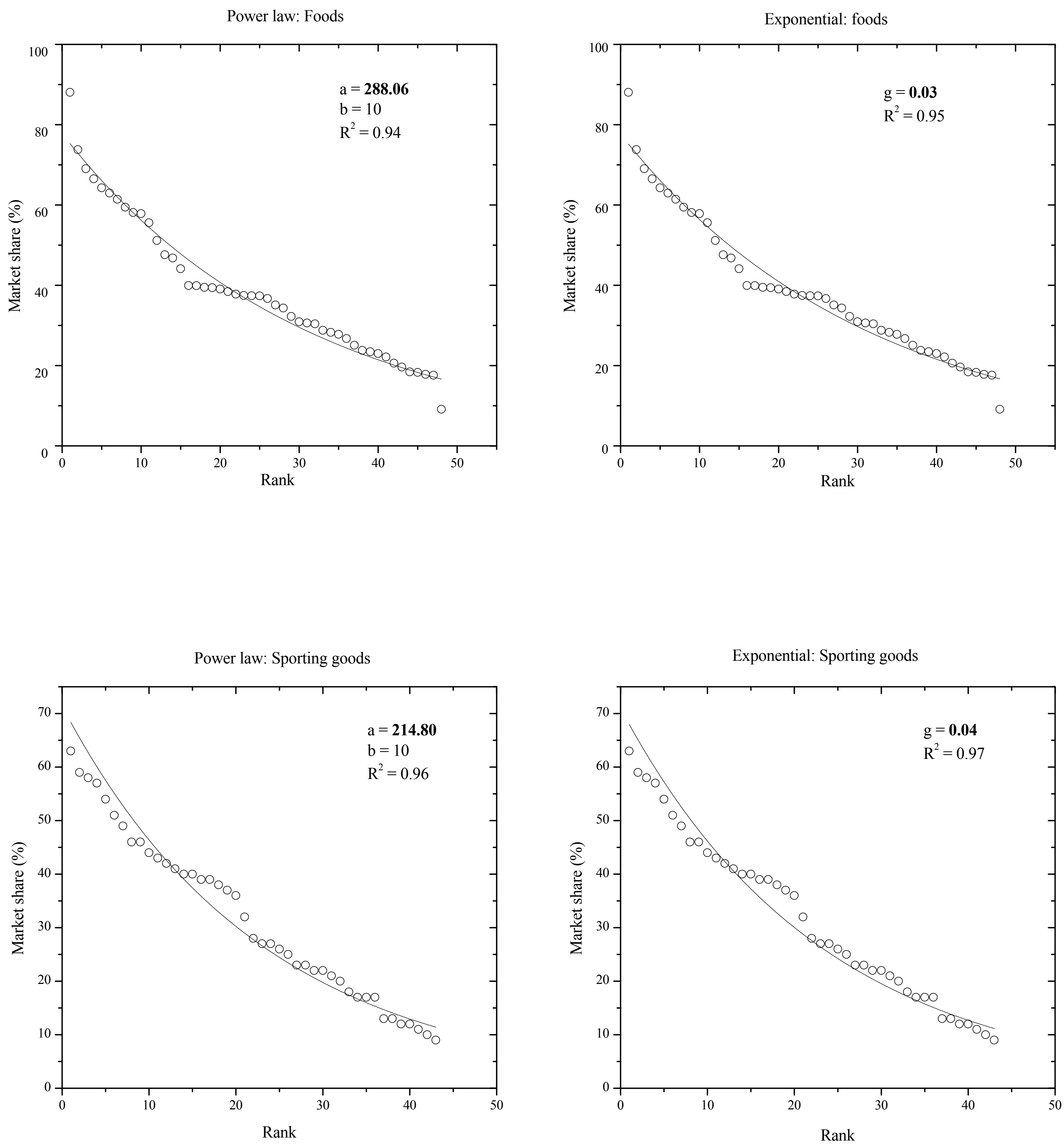


FIGURE 5: Averages of market shares with same rank, across product categories of foods