BEBR

FACULTY WORKING
PAPER NO. 1004

A Game Theoretic Rationale for Henderson's Rule of Three and Four

K. Ravi Kumar
Howard Thomas

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 1004

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

January 1984

A Game Theoretic Rationale for
Henderson's Rule of Three and Four

K. Ravi Kumar, Professor
Department of Business Administration

Howard Thomas, Professor
Department of Business Administration

Abstract

Henderson's hypothesis of the rule of three and four (Henderson, 1979: 90) conjectures that a stable competitive market never has more than three significant competitors. Henderson observes that the rule appears to be a good prediction of the results of competition in such fields as steam turbines, automobiles, baby food, soft drinks, and airplanes.

A game theoretic model is formulated which suggests that aggregation of competitive information and imposition of structural stability results in an explanation of the rule in a generalized context.

Introduction

It has been acknowledged (Cunningham and Robertson (1983)) that certain consulting firms and consultants have contributed in large measure to the development of the theory of strategic management. Henderson (1979), the founder of the Boston Consulting Group, has been active in the formulation of both theory and practice in strategic management and has contributed many strategic perspectives based mainly upon his experience and observation of many strategic decision situations.

Henderson discusses the anatomy of competition in a number of important contributions (1979: 90), (1983: 7). He stresses the value of understanding the nature of competitive equilibrium since it provides a benchmark for the formulation and evaluation of competitive strategy.

In this paper one of Henderson's propositions about competition and corporate management is examined. Henderson states the Rule of Three and Four as follows (1979: 90):

> A stable competitive market never has more than three significant competitors, the largest of which has no more than four times the market share of the smallest.

Henderson's rule is based upon empirical observation and appears to fit the results of effective competition in a range of industries.

Henderson (1979: 93) later points out that "a rigorous application of the Rule of Three and Four would require identification of discrete, homogeneous market sectors in which all competitors are congruent in their competition."

This latter statement could be interpreted in two ways. First that the three or four significant competitors adopt different strategies. Second, that the three or four significant competitors represent the three or four viable competitive positions within the industry. These positions are essentially the basis for the formulation of strategic groups (Caves and Porter (1977), Porter (1980)).

The main aim of this paper is the development of a game theoretic model which presents an explanation of the Rule of Three and Four in a generalized context. The implications of the results are discussed and some research uses of game theory in identifying strategic groups are then suggested.

Are There Limits to the Number of Strategic Groups

Strategic groups are defined to be groups of firms that possess or use similar strategies, i.e., goals and objectives, product-market competitive positioning, resource allocations and operating policies. These groups could be within an industry or across industries (Cooper and Schendel (1971)).

Consider groupings within an industry, say, the beer industry. If we consider the marketing strategy, each producer tries to differentiate his own product from the rest of the competitors. For example, Miller High Life is directed at hard-working blue-collar workers for relaxation at the end of a tough day; Michelob for those weekends with close friends. In this instance, the number of strategic groups is equivalent to the amount of product differentiation possible in that industry. Levitt (1980) argues that all goods and services are differentiable--

even if the generic product is identical, the offered product is differentiated. Even though the above argument does not explain why the product should be differentiated, but rather than it can be, it does imply that differences in consumers and/or producers may necessitate product differentiation.

The question then remains whether there is any limit to the number of strategic groups. The implication in Levitt (1980) is that, potentially, there is none; "only the budget and the imagination limit the possibilities." On the other hand, oligopolistic forms of market structures are observed for many stable products, such as automobiles, steel, toothpaste, etc. Often three or four firms absorb most of the market share in industries such as turbine-generators (Sultan (1975)). Schendel and Patton (1978), in their analysis of the beer industry, choose three hypothesized groupings based on geographic scope, namely small regional, large regional and national. Aggregating all the firms into an industry level model produced severe heterogeneity whereas the geographic grouping was hypothesized to be homogeneous within groups. Indeed, these persistent differences in industry structure lead to relatively few market segments suggesting that there might be limits to the number of strategic groups.

In the following section, a game theoretic model is formulated which examines whether there are, in fact, only a limited number of strategic positions that firms can occupy in an industry and considers the stability of these positions.

A Game Theoretic Formulation

The model assumes that the industry consists of a large number of firms, each of which produces a single product for the market. Each firm's product can be differentiated, but it remains a close substitute for the products of other firms. The firms act as monopolists, as far as their customers are concerned, and set prices; on the other hand, they compete amongst themselves for the consumer demand since customers can switch from one firm's product to another's. This economic model was posited by Chamberlin to capture the notion of product differentiation; this market structure falls in between that of pure competition and pure monopoly, and is aptly named monopolistic competition (Mansfield, Chap. 24 (1974)). The basic elements in the modelling approach of individual firms used here will be goals/objectives, means or resource decisions possible and environmental constraints (Schendel and Patton (1978)).

The model assumes that there is a single objective for each firm, namely utility maximization. While it is true that there very well could be a number of goals, often conflicting, that firms operate under, the contention here is that cost factors could be attributed to all the key goals and a net utility function could be derived for each firm. Essentially, the trade-off between, say profit maximization and market share maximization, can be captured in the utility function.

The actions or resource decisions, that each firm can take, are made over several controllable variables. These can be strategic and operating variables. For ease of exposition, only strategic variables will be allowed to influence the utility function. One of these will be the price (P_i) that the firm i charges for its product, thus acting as a monopolist towards its consumers. Two other strategic variables that will be used are advertising (A_i) and material cost (M_i). These variables were chosen from the table of variables listed for the brewing industry by Schendel and Patton (1978).

The environmental constraints encompass those variables that the individual firm cannot control even though its decisions may influence them. These non-controllable variables will be chosen to describe the competitive nature of the industry. Here, the non-controllable variable used, for exposition's sake, will be the industry average price ($\overline{P}$), the industry average advertising ($\overline{A}$) and the industry average material costs ($\overline{M}$). It should be noted that these variables are the industrial counterpart to the firm controllable variables. Each firm takes these variables as given while trying to maximize their utility function.

Then, in this model, each firm, assumed to be identical, will maximize its utility function

$$U(P_i, A_i, M_i, \overline{P}, \overline{A}, \overline{M}).$$

The function U is assumed to be continuously differentiable in all its variables but not necessarily quasi-concave in its strategic variables (P_i, A_i, M_i). An equilibrium for this game is a set of strategies for each firm such that the outcome of the game is stable, i.e., in equilibrium, firm i cannot do any better than using its equilibrium strategy given that its competitors are playing their equilibrium strategies. This type of stability concept is called a Cournot-Nash equilibrium in economics literature. In particular, symmetric equilibria will be examined since a priori all firms have been assumed to be identical.[1] See the appendix for a more formal definition.

The symmetric equilibrium could be a pure strategy equilibrium in which case all firms are following the same strategy. Then, there is only one strategic group consisting of all firms in this equilibrium. On the other hand, the symmetric equilibrium can be in mixed strategies, i.e., equilibrium dictates that each firm randomize over some distinct number of strategies. It is not reasonable nor practical to expect firms to toss coins in making strategic decisions (and also constantly changing them) in reality. Then the interpretation of these mixed strategy equilibria will be that the firms, in the industry, split into different groups, each group using a distinct strategy specified by the mixed equilibrium. The proportion of firms in each group will approximate the probability assessment assigned to the particular strategy used by the group in the equilibrium.

Intuitively, given a symmetric mixed equilibrium, the firms are indifferent to the actual choice between any of the strategies in the equilibrium since the utility level associated with each is the same.[2] The actual allocation of firms to these strategies is not possible at this level of generality but it is presumed that the prior history of

---

[1] It should be noted that there could very well be asymmetric equilibria for a symmetric game.

[2] It is possible for other strategies to have a similar utility level if the probability associated with these are zero. It will be assumed that this does not occur, i.e., all the strategies producing maximal utilities have positive probabilities associated with them.

the firm will make it gravitate to a specific strategic position. The point of focus in this research is the question of how large the number of strategic positions can be in this equilibrium.

Limits to the Number of Strategic Groups

A priori, there is no reason to believe that there is a bound on the number of strategic positions for the model presented. However, by imposing the condition of structural stability in the equilibrium structure, a definite limit on strategic positions emerges.

The notion of Cournot-Nash equilibrium intimated a sense of stability of the following form: given the parameters of the model, once every firm is playing its equilibrium strategy, then no firm individually has any incentive to break away from this equilibrium. The idea of structural stability argues for yet another sense of stability, namely: given small shifts in the parameters of the model, Cournot-Nash equilibrium structure changes only slightly. This sort of stability ensures that small random fluctuations of the environment will not make firms change their strategies drastically.

This concept of structural stability is taken from the field of differential topology in mathematics (Guillemin and Pollack (1974), Chapters 1 and 2). Its relevance to the model presented here is obvious since, in practice, the utility function, which is the fundamental building block, can only be estimated within some non-zero margin of error. Given small perturbations in the utility function, it is important to determine whether the equilibrium, in these perturbed models, is similar and close to that in the unperturbed model.

If so, the equilibrium in the unperturbed model is said to be structurally stable.

To understand how structural stability is maintained, the equilibrium structure should be examined a little more closely. In the model, given industry information $(\overline{P}, \overline{A}, \overline{M})$, each firm chooses values for its strategic variables (P_i, A_i, M_i) so as to maximize its utility. Assume, for the sake of exposition, that the mixed equilibrium has exactly two strategies (P_1^*, A_1^*, M_1^*) and (P_2^*, A_2^*, M_2^*) with the equilibrium industry variables being $(\overline{P}*, \overline{A}*, \overline{M}*)$. Then, due to their optimality, the utility associated with (P_1^*, A_1^*, M_1^*) and (P_2^*, A_2^*, M_2^*) given $(\overline{P}*, \overline{A}*, \overline{M}*)$ is exactly the same or, to put it another way, the difference in the utility functions, at the equilibrium positions, is zero.

Now perturb the utility function, structurally, by a very small amount so as to keep the values of the functions and its derivatives close to their originally unperturbed values. To ensure stability of the unperturbed equilibrium, the utility-difference map has to be checked to ensure that it is well defined and that it assumes a value of zero close to the unperturbed equilibrium industry information value of $(\overline{P}*, \overline{A}*, \overline{M}*)$.[3]

The model assumptions assure that the utility difference map is well defined but it is the existence of a value of zero for this map near $(\overline{P}*, \overline{A}*, \overline{M}*)$ which brings about limitations on the number of

---

[3]The explicit conditions are given in the appendix with mathematical rigor.

strategic groups. Note that in this example equilibrium, the utility difference map is a function of the industry information $(\overline{P}, \overline{A}, \overline{M})$ and the value it takes is the difference between two strategic positions. Hence it takes on a single value, i.e., the domain of this map is the space of values that $(\overline{P}, \overline{A}, \overline{M})$ can take, i.e., R^3 and the range of the map is the real line R^1.

Now suppose there were, in fact, five strategic positions in equilibrium. Then, the utility difference map would have the same domain as before but its range would be values in the four dimensional real space R^4, i.e., it would have to specify values of utility differences between the first and second strategic positions, the second and third, the third and fourth and finally the fourth and fifth[4]--four different values for every value of $(\overline{P}, \overline{A}, \overline{M})$. For the original equilibrium $(\overline{P}^*, \overline{A}^*, \overline{M}^*)$ to be structurally stable, this perturbed utility difference map has to take on a value equal to zero (in the four dimensional space) near $(\overline{P}^*, \overline{A}^*, \overline{M}^*)$.

In general, if a map is from a higher dimensional to a lower dimensional space (R^3 to R^1 as in the first example), then each value that this map takes on can be reproduced for some specific perturbation of this map. <u>But</u>, when this map is from a lower dimensional to a higher dimensional space (R^3 to R^4 as in the second example), there exists some perturbation that misses any and every specified value that the unperturbed map takes on. In the present model, the value of the utility difference map that is of interest is zero. Mathematically,

---

[4]There is no need to specify the difference between the first and fifth since it will become redundant.

for a two strategy equilibrium, perturbations could take on this value, but for a five strategy equilibrium, there exists a perturbation which never takes the value zero. Hence the second example can never be structurally stable while the first possibly could be.

An intuitive example of this mathematical result is the following: consider the intersection of straight lines--two lines can intersect at a point and small perturbations of these lines will still intersect at a point. On the other hand, let three lines intersect at a point--then there is always a small perturbation of these lines such that these lines do not intersect at a common point. This example shows that two lines intersecting at a common point is stable but anything above two is unstable, i.e., the nature of this intersection changes drastically.

This result is stated below as:

Instability Theorem: If the dimension of the industry information space (or the uncontrollable variable space) $(\overline{P}, \overline{A}, \overline{M})$ is three, then no mixed symmetric Cournot-Nash equilibrium with more than four strategies can be structurally stable.

A rigorous proof of this result is given in the appendix.

Discussion of Result

The result implies that if there are three industry information variables $(\overline{P}, \overline{A}, \overline{M})$ involved, then the firms cannot split up into more than four strategic groups, each group differentiating the product uniquely, in a structurally stable equilibrium. Notice that the result does not necessarily imply the existence of a stable equilibrium with exactly four groups. It is possible to have a product differentiated stable equilibrium with only two groups of firms. Also, an example could be constructed, within the framework of this model, which involves more than four groups in equilibrium. However, any perturbation of the example would destroy the structure of this equilibrium. Kumar (1981, Chapter 2) has examples of such equilibria derived from consumer behavior in an economics context. This substantiates the existence of such equilibria in realistic models and also that they are not degenerate.

Similar dimensional arguments have been used, with inverted logic, in the field of multi-dimensional scaling (Churchill (1976), pp. 233-241). The argument is that given n firms in a product differentiated market, they can be represented in a (n-1) dimensional attribute space with no constraints on the relationships between these firms. In a sense, this reflects the idea that given an industry with four strategy groups, one can find three factors over which they can be clearly differentiated and still retain the independence of each factor. This is similar to the result here which claims that given three factors ($\overline{P}$, $\overline{A}$, $\overline{M}$) there can be, at most, four strategic groups in a stable equilibrium.

Given the instability theorem for this model, a clearly important problem is how to define the characteristics of the four possible strategic groups that can occur in a maximally differentiated equilibrium. This issue, though not formally treated here, can be handled using the well-known concepts of efficient frontiers and stability.

Summary and Conclusions

The game theoretic formulation and proof of the validity of Henderson's conjectures depends upon a number of assumptions made in the analysis. These are as follows: first, that there are a large number of competitors in the industry at the outset so that the aggregate statistics that each competitor faces are the same; second, the notion of stability in the process of industry evolution.

The idea of stability and equilibrium is important from a strategic management viewpoint. If a particular company in an industry can predict the equilibrium strategy positions when the industry matures, then it can adapt its strategy to position itself within one of the stable strategy positions and thus not drop by the wayside in any subsequent 'shakeout'. Alternatively, it may attempt to extricate itself from an unfavorable strategy position assuming an accurate prediction of the positioning of the long-term equilibrium.

The proof in the paper, albeit based upon a simplified model, tends to confirm Henderson's empirical observations. The vision of mature industries given by this model is that there are only a few pure strategies being used by firms in the industry--either there are a few main firms using different strategies or there are many firms "bunching up" or grouping around a few distinct strategies. The former is the rule of 3 or 4 by BCG and the latter is a generalization of the Rule. The generalization probably narrows to the Rule of 3 and 4 when the fixed costs to entry are very large.

References

Caves, R. E. and M. E. Porter (1977), "From Entry Barriers to Mobility Barriers: Conjectural Decisions and Contrived Deterrence to New Competition," Quarterly Journal of Economics, 91, May, pp. 241-262.

Churchill, G. A., Jr. (1976), Marketing Research Methodological Foundations, Dryden Press, pp. 233-241.

Cooper, A. C. and D. E. Schendel (1971), "Strategy Determination in Manufacturing Firms: Concepts and Research Findings," Proc. Amer. Marketing Assoc., Fall Conference, Minneapolis.

Cunningham, W. H. and T. S. Robertson (1983), "Editorial to Marketing Strategy Special Issue," Journal of Marketing, Spring, p. 6.

Guillemin, V. and A. Pollack (1974), Differential Topology, Englewood Cliffs, New Jersey, Prentice-Hall, Inc.

Henderson, Bruce (1979), Henderson on Corporate Strategy, Cambridge, Mass.: Abt Books.

Henderson, Bruce (1983), "The Anatomy of Competition," Journal of Marketing, Spring, pp. 7-11.

Kumar, K. Ravi (1981), "Product Differentiated Equilibrium: Structural Stability, Maximal Dispersion and Realization," unpublished Ph.D. dissertation, Northwestern University.

Levitt, Theodore (1980), "Marketing Success Through Differentiation--Of Anything," Harvard Business Review, Jan.-Feb., pp. 83-91.

Mansfield, E. (1974), Economics: Principles, Problems, Decisions, New York, W. W. Norton and Company.

Porter, M. E. (1980), Competitive Strategy, New York, The Free Press.

Schendel, D. E. and G. R. Patton (1978), "A Simultaneous Equation Model of Corporate Strategy," Management Science, Vol. 24, No. 15, pp. 1611-1621.

Sultan, R. G. M. (1975), Pricing in the Electrical Oligopoly, Vol. II, Harvard University Press, Cambridge, Mass., Chapter 2.

D/165

Appendix

For the given model, the strategic variables for the firm are (P_i, A_i, M_i), the industry information variables are $(\overline{P}, \overline{A}, \overline{M})$ and the common utility function facing each firm is

$$U(P_i, A_i, M_i, \overline{P}, \overline{A}, \overline{M})$$

which is continuously differentiable. Each firm maximizes its utility function over its strategic variables taking the industry information variables as given.

Definition: A mixed strategy based on k pure strategies is given by $((P_i, A_i, M_i, \omega_i)_{i=1}^k)$ where ω_i is the probability placed on the i^{th} strategy (P_i, A_i, M_i) with $\omega = (\omega_1, \ldots, \omega_k) \varepsilon \Omega_k \equiv \{(\mu_1, \ldots, \mu_k \mid \sum_{i=1}^k \mu_i=1\}$.

Definition: The reaction set at $(\overline{P}, \overline{A}, \overline{M})$, denoted by $R(\overline{P}, \overline{A}, \overline{M})$, is the set of (P_i, A_i, M_i) which globally maximizes the utility function at $(\overline{P}, \overline{A}, \overline{M})$.

Definition: A symmetric mixed strategy Cournot-Nash equilibrium based on k pure strategies is given by $((P_i^*, A_i^*, M_i^*, \omega_i^*)_{i=1}^k, \overline{P}^*, \overline{A}^*, \overline{M}^*)$ satisfying for all $\hat{\omega} = (\hat{\omega}_1, \ldots, \hat{\omega}_k) \varepsilon \Omega_k$.

$$\sum_{i=1}^k \omega_i^* U(P_i^*, A_i^*, M_i^*, \overline{P}^*, \overline{A}^*, \overline{M}^*) \geq \sum_{i=1}^k \hat{\omega}_i U(P_i^*, A_i^*, M_i^*, \overline{P}^*, \overline{A}^*, \overline{M}^*),$$

$$\sum_{i=1}^k \omega_i^* P_i^* = \overline{P}^*, \quad \sum_{i=1}^k \omega_i^* A_i^* = \overline{A}^*, \quad \sum_{i=1}^k \omega_i^* M_i^* = \overline{M}^*,$$

$$\text{where } \omega^* = (\omega_1^*, \ldots, \omega_k^*) \varepsilon \Omega_k, \ (P_i^*, A_i^*, M_i^*) \varepsilon R(\overline{P}^*, \overline{A}^*, \overline{M}^*)$$

The above definition implies that there are k strategies $(P_i^*, A_i^*, M_i^*)_{i=1}^k$ which are the best responses given $(\overline{P}*, \overline{A}*, \overline{M}*)$ and that each of these strategies has a corresponding probability ω_i^* associated with it. This probability $\omega^* = (\omega_1^*, \ldots, \omega_k^*)$ is the one that gives the highest expected profits compared to all other probability measures and also satisfies the consistency condition that the expected or average value of the best response strategy is indeed $(\overline{P}*, \overline{A}*, \overline{M}*)$.

This equilibrium implies that around a small neighborhood of $(\overline{P}*, \overline{A}*, \overline{M}*)$, there exist k local maxima and by the assumption of positive probabilities associated with each utility maximizing strategy (see footnote 2), they are isolated. Also, the utility levels at each of these local maxima, at $(\overline{P}*, \overline{A}*, \overline{M}*)$, is the same. Using notations,

let $x_i \equiv (P_i, A_i, M_i)$, $i=1, \ldots, k$

and $y \equiv (\overline{P}, \overline{A}, \overline{M})$

and $D: R^3 \to R^{k-1}$ defined by

$$D(y) = (U(x_i(y),y) - U(x_{i+1}(y),y), \quad i=1,2, \ldots, k-1, \quad x_j(y) \in S(y), \quad j=1, \ldots, k)$$

where S(y) is the set of local maxima at the point $y = (\overline{P}, \overline{A}, \overline{M})$. Then, in a small neighborhood of $y* = (\overline{P}*, \overline{A}*, \overline{M}*)$, D(y) is well defined and $D(y*) = 0$.

Now consider all perturbations U_λ of the function U such that the values of the function U_λ, its first and second derivatives can be made as close as needed to those corresponding values of U by choosing λ small enough. To ensure the stability of the equilibrium using profit function U, it is necessary to demonstrate the existence of a similar equilibrium, close by and unique in a neighborhood of the original equilibrium, for all perturbations U_λ, with λ in a some open neighborhood of the value 0 (where it is assumed that $U_0 = U$). Three conditions have to be satisfied for such an existence, namely

(1) the local maxima sets S(y) and $S_\lambda(y)$ must be close to each other in a neighborhood of y* and the cardinality of $S_\lambda(y)$ should be equal to k.

(2) if condition (1) is satisfied, then the profit difference function D_λ is well defined in this neighborhood of y*. Then, it is necessary to show the existence of $y_\lambda^* \equiv (\overline{P}_\lambda^*, \overline{A}_\lambda^*, \overline{M}_\lambda^*)$ close to y*, in the neighborhood were D_λ is defined, such that $D_\lambda(y_\lambda^*) = 0$.

(3) Then it is necessary to ensure the existence of a probability vector ω_λ^* such that the expected value of the best response strategy $(P_{i\lambda}^*, A_{i\lambda}^*, M_{i\lambda}^*)$ is indeed $(\overline{P}_\lambda^*, \overline{A}_\lambda^*, \overline{M}_\lambda^*)$.

Condition (1) can be shown to be true through an application of the implicit function theorem and using the fact that the local maxima are non-degenerate critical points (Guillemin and Pollack (1974)). It is the second condition which leads us to the limitation on the value that k can take.

Proof of the Instability Theorem:

Consider the case where $k > 5$. Then, the continuous map D, whose domain is some compact subset A of the neighborhood of y^* and whose range is a subset B of R^{k-1}, satisfies $D(y^*) = 0$. It should be noted that D is uniformly continuous and B is a compact set of measure zero in R^{k-1}. Now either 0 is on the boundary of set B or not.

If 0 is on the boundary, then there exists a sufficiently small vector translation D_λ such that 0 is not contained in the range of this perturbation D_λ.

If 0 is in the interior of B, then approximate D (which is uniformly continuous) uniformly by means of a piecewise continuous linear function, $\hat{D}$. Note that this is a perturbation of D. For this approximation, 0 is either in its range or not. If 0 is not in its range, then this perturbation suffices and let $D_\lambda = \hat{D}$. If 0 is in its range, then it must be on its boundary (since $\hat{D}$ is piecewise linear) and therefore a sufficiently small vector translation D_λ of $\hat{D}$ can get rid of the 0 in its range.

Thus, we can construct a perturbation D_λ of D either through pure vector translation or through a uniform approximation or through a combination of both, such that no y exists with $D_\lambda(y) = 0$.

$\square$