LACEA 2001
MONTEVIDEO URUGUAY

LACEA 2001
MONTEVIDEO URUGUAY

Salvador Valdés - Prieto

Optimal Price and Marketing Policies with Stock Demand

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
March 28, 2001 **Comments welcome**

Optimal price and marketing policies with stock demand*

by
Salvador Valdés-Prieto**

Abstract

This paper presents a model of optimal pricing and marketing policies in "stock demand" markets, such as the financial services industries, cell phones and mandatory "privatized" social security. In these markets the clientele with contracts evolves slowly. It is shown that in a perfect equilibrium, the unique optimal price path leads to a steady state where the effective price elasticity depends on the rate of depreciation of the clientele through exit, on the discount rate, and on the elasticity of both entry and exit flows to prices. A perfect symmetric equilibrium does not exist always, however. Under some parameter conditions there can be multiple asymmetric Nash equilibrium in marketing investment, which are more reasonable than the symmetric equilibrium because the latter does not meet a refinement we propose. The asymmetric equilibria lead to a war of attrittion and to mergers that concentrate the industry. It is shown that in the case of mandatory demand, for an important region of parameter values, only concentrated industry structures allow a perfect symmetric equilibrium to exist, while excessively atomistic structures encourage marketing wars and mergers.

JEL classification numbers: D4, L1, M3

---

* This work received support from the Research Fund of the School of the Economics and Business Administration of the Catholic University of Chile. I am grateful for the comments of Jorje Tarziján, Edgardo Barandarián, Ronald Fischer and other participants at a seminar at Baños del Corazón, January 22-24, 2001. The comments of Felipe Zurita helped substantially. Of course, the only one responsible for errors and ommissions is the author.

** Full Professor at the Instituto de Economía, Pontificia Universidad Católica de Chile.

1. Introduction

The economic analysis of competition in any industry requires a model of price and marketing decisions. The conventional model treats demand as a flow, which may be influenced by prices and by the commercial messages issued by the marketing investment made by the firm. Examples of important recent studies that have used this approach are Roberts and Samuelson (1988) and Rizzo (1999).

This paper explores the argument that the conventional model is not satisfactory for industries where clients confront significant costs when they switch among providers on a high frequency basis. In a number of important industries, such as cell phone subscriptions, magazine and newspaper subscriptions, the financial services industry (mutual funds, life insurance, pension fund management), education and health insurance, the client obtains services on a continuous basis from the same provider, until he or she switches provider, an event that happens rather infrequently.

Continuous service could be understood as continuous choice of provider in a frinctionless world. However, the gain from getting service from the best rather than the current provider for a single day (or any short period) is swamped by the fixed transaction and computation costs of identifying the best provider for the day. Thus, transaction costs make it suboptimal for the consumer to change provider on a daily basis.

It is natural for a firm confronted with this type of client to make a distinction between the clientele that has already signed a service contract that allows them to claim service on a continuous basis, and other potential clients that have not signed up. Although the members of the former group may be free to end their contract when they wish to, in practice the switching cost they face assures that their withdrawal occurs “slowly” over time. This fact may justify treatment of the clientele that has a service contract as a “stock demand”, which evolves slowly over time, and not as a flow that responds instantaneously as in conventional models. In general, both type of clientele may coexist. For example, newspapers are sold both on a subscription basis and on newstands, and prices are different.

This paper presents a model of optimal uniform pricing for stock demand. It is shown that the unique optimal price path leads to a steady state where the effective price elasticity depends on the rate of depreciation of the clientele through exit, and on the elasticity of both entry and exit flows to prices. The discount rate used by the firm has a direct influence on pricing. The exit rate can be interpreted as a “churning” rate that is optimized. Separate marketing investments are allowed both to attract new clients and to improve the quality of service, and optimal marketing policies are identified.

One alternative modelling strategy is to propose a distinction between a “long-run” demand and a “short-run” demand, linked through a partial adjustment mechanism[1]. Partial adjustment assumes an exogenous and fixed coefficient of adjustment. Instead, the stock demand approach optimizes the complete policy path, emphasizing the transition path in addition to the steady state, and linking them optimally. Near the steady state of the stock demand, the optimal path can be linearized and takes a form similar to partial adjustment, with the difference that the coefficient of adjustment is chosen by the firm to maximize profits[2]. As the stock demand model is better founded on microeconomic analysis, it is more appealing than the partial adjustment model.

An important empirical regularity across consumer markets is that concentration remains above some threshold for large ranges in the size of the market, as documented and justified by Sutton (1991). We find that a perfect symmetric equilibrium does not exists always in a stock demand model. Under some parameter conditions there can be multiple asymmetric Nash equilibrium in marketing investment, which are more reasonable than the symmetric equilibrium because the latter does not meet a refinement we propose. The asymmetric equilibria lead to mergers that concentrate the industry.

It is shown that in the case of mandatory demand, for an important region of parameter values, only concentrated industry structures allow a perfect symmetric equilibrium to exist, while excessively atomistic structures encourage mergers. For other ranges of parameter values, the merger process continues until a single firm is left in the industry. This case is interesting, as a growing number of industries face mandatory demand.

---

[1] For example, $D^{shortrun} = A(t) = A(t-1) + \alpha[D^{longrun}(p(t),M(t),m(t)) - A(t)]$, with α = adjustment coefficient in [0,1], exogenous.

[2] Another difference arises when a given (say rising) price path is chosen by the firm: in the partial adjustment model, this path shifts the long run demand over time keeping the speed of adjustment constant, while in the stock demand model, the target steady state remains constant over time and the speed of adjustment is varied optimally.

Section 2 proposes a simple model for decisions of a firm facing a stock demand, under uniform pricing . Section 3 discusses the first order conditions for the steady state policies for a single firm, emphasizing how different they are from those that arise from the flow demand approach. Section 4 studies the Nash equilibrium in the subgame that starts in the steady state and determines the range of parameter values under which the “reasonable equilibria” are symmetric or assymmetric. Section 5 concludes.

2. Decisions under stock demand and uniform pricing

The conventional models of a firm with flow demand take the following form in their simpler version[3]:

(1) $$\underset{\{p(t);M(t)\}}{Max}\left\{V = \int_{t=0}^{\infty}[p \cdot D(p,M) - F - c_a \cdot D(p,M) - c_m \cdot M]\, e^{-rt}dt\right\}$$

In ecuation (1), the firm maximizes the present value V of its profits, chosing the level of a uniform price p and the number of messages M. M can also be interpreted as an index of quality. Both affect the flow demand D(p,M). In addition to a fixed cost F that accounts for economies of scale both in the back office and in marketing (say, sales-force training), expenditures include the variable production cost $c_a \cdot D(p)$ and marketing investment $c_m \cdot M$. The latter two are simplified by assuming constant returns to scale[4]. It is well known that que the optimal pricing policy meets in every instant the Lerner Rule ($(p^*-c_a)/p^* = 1/\eta_{D,p}$), and that the optimal marketing investment policy meets in every instant

[3] It is easy to extend these models to consider a case where demand depends of the stock S of marketing investment, so D = D(p,S), where S evolves according to $dS/dt = M - \delta \cdot S$, where δ = depreciation rate of the marketing stock, in the mind of consumers.

[4] In the case of education it is clear that c_a rises above some size of school, so this aspect of the model should be generalized for an application to education.

the equation $(p-c_a)\cdot\partial D/\partial M = c_m$. The discount rate r does not intervene in these decisions. See Berndt (1991) for a summary.

The model we present in this section has the same cost structure as conventional model (1), but the demand is of a "stock" nature, as explained in the Introduction. Further, we assume that the firm is free to change its price p over time, provided the price is uniform for all clients. This set up is standard in the mutual fund management market[5].The model we consider is:

$$(2)\qquad \underset{\{p(t);M(t)\}}{Max}\left\{V=\int_{t=0}^{\infty}[p\cdot A-F-c_a\cdot A-c_m\cdot M-k_m\cdot m\cdot A]e^{-rt}dt\right\}$$

$$subject\ to:\quad 1)\ \frac{dA}{dt}=E(p,M,m)-f(p,m)\cdot A$$

$$2)\ \ 0\le A(t)$$

$$3)\ \ A(t)\le A^U$$

$$4)\ \ M(t)\ge 0; m(t)\ge 0$$

$$and\ the\ initial\ condition\qquad A(0)=A_o$$

where A_t = stock of clients, i.e. members of the clientele, as of date t.

p_t = the uniform price charged in period t ($/member).

M_t = number of marketing messages issued in t to attract new members, at a marginal cost of c_m each. These may be called pre-sale messages.

m_t = number of marketing messages issued in t to "maintain" current members, expressed as a rate per member, at a marginal cost of k_m each. These may be called post-sale messages and are distributed uniformly to all members. Alternatively, m can be interpreted as an index of quality of service. In financial firms, the main dimension of service quality is financial performance, i.e. the expected return and risk earned by the funds owned by clients and managed by the firm.

E_t = positive transfers (entry of new members).

---

[5] In the U.S., the Investment Company Act of 1940 requires mutual funds to apply price uniformity regarding expenses and management fees. Firms whose prices are set by contract are usually free form this regulation, but there is a number of cases where prices are set by contract for a fixed period, and at its expiration, the firm offers the client another contract. These other cases may require adjustments to model (2).

A^U = universe of potential members, which may be the overall market size.

ϕ_t = exit *rate* among current members, as of date t.

$S_t = \phi(p,m)\cdot A$ = negative transfers (exit of existing members).

The first restriction in program (2) indicates the stock of members evolves slowly, in response to entries E and exits S. Entries E depend negatively on the uniform price ($\partial E/\partial p < 0$) and positively of marketing investment in attractive messages ($\partial E/\partial M > 0$). If the current investment in maintenance messages is taken by clients as a signal of the level that such messages will attain in the future, and if clients value these messages, entries may also depend positively on the current investment in maintenance messages ($\partial E/\partial m > 0$). Exits S depend on the stock of members served by the firm (A). In addition, the exit rate ϕ depends positively on the level of firm's price ($\partial\phi/\partial p > 0$) and negatively of marketing investment in "maintenance" messages, expressed as a rate per member ($\partial\phi/\partial m < 0$). Note that the rivals' attractive marketing enters the exit rate as well, but it is not shown because it is not under the control of the firm.

The functional form for exits is different from the one for entries because the former depends on the stock of clients, A. This model emphasizes that exits should fall proportionately when the stock of members is smaller, when the other controls are kept constant. By symmetry, the inverse phenomenon for entries would be that entries should fall when the stock of members *not* served by the firms rises. For example, when the set of potential members that are not served by this firm (A^U-A_t) falls, then entries should also fall proportionately, provided the other controls are kept constant. When the firm has a small market share, the effect of (A^U-A_t) becomes very small. When this effect remains near zero it is effectively a constant, so it can be safely ommitted. Thus, (2) represents the problem for a firm with a small market share.

In practical problems, identifying messages that attract new clients may require some effort. For example, if a mutual fund manager pays a salesperson a commission in proportion to increments in funds attracted from his or her client list, then the salesperson may see an incentive to churn the funds of its existing client list, as this increases the increments in funds - at the expense of increasing decrements as well - with little attraction of new clients. But when a mutual fund manager pays commission on the basis of the number of new clients and their funds, this investment may be represented safely by M.

The second restriction in (2) indicates that the stock of clients cannot be negative. The third restriction indicates that the stock of clients cannot surpass some constant A^U which is the total number of clients in all firms in the industry. This constant can represent the potential market of registered clients[6], or a legal restriction to market share[7], or a restriction to market share imposed by a cartel to its members. Model (2) represents the case of a relatively small firm, defined as one whose number of members is far below the ceiling A^U. For this reason, in what follows we ignore the third restriction and the influence of the remaining market (A^U-A) over the flow of entries[8]. The fourth restriction represents the fact that marketing investment is sunk, or irreversible. It prevents the firm from earning income by buying commercial messages from clients and reselling them to salespeople.

We have ommitted from program (2) other arguments of demand, such as the prices charged by rivals, the messages issued by rivals, and the quality of service of rivals. This ommission requires an explanation, because if rivals are modifying their policies over time, the entry and exit functions should include explicitly the time argument. And given that optimal policy is not static for any given firm, it is likely that rivals are indeed changing their policies over time. Of course, this would not be true in a steady state, a point we exploit.

The general game the firm faces is an oligopoly where N firms choose sequences of actions over time. Rather than attempting to find the equilibria of this game taking into account its full complexity, we simply ask if there exists an equilibrium where the actions of each firm are constant over time, i.e. if there is any equilibrium that is a steady state. A candidate steady state equilibrium should meet at least two conditions:

(a) When the rivals take the actions prescribed by a possible steady state equilibrium, it must be the case that the firm's best response converges to a steady state, so it also chooses actions that remain constant over time. Conversely, if the firm's actions do

---

[6] In a mandatory market, A^U is given by the coverage of the mandate.

[7] For example, in the new Mexican pension system with individual accounts created in 1997, no firm is allowed to serve more than 20% of potential members.

not converge to a steady state when other firm's actions remain constant, then no steady state equilibrium exists. This question can be answered using program (2). The Appendix shows that the optimal actions of a single firm do converge to a steady state, as shown in figure 1, and as described in section 3 in more detail.

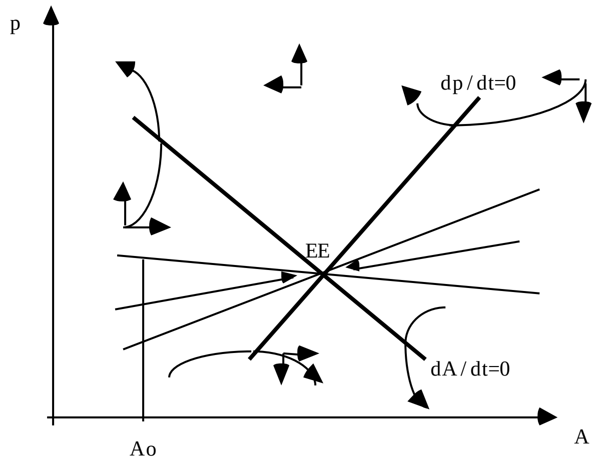

Figure 1: Phase diagram for problem (2)

(b) Subgame perfection. The actions prescribed by the steady state of any firm must be part of a Nash equilibrium between the firms in the subgame where all firms can change actions simultaneously. If this were not the case, then at least one rival would wish to deviate from the candidate steady state, and no steady state equilibrium would exist. Then program (2) would be an inappropriate representation of the firm's decision problem, because rivals would be changing their actions over time, and the entry and exit functions in (2) should include explicitly the time argument. Alternatively, the assumed constancy of rivals' actions would not be credible. The existence and uniqueness of a Nash equilibrium in this game are discussed in section 4 below.

3. Optimal policies for a single firm

[8] We lift this assumption in section 3.4.

This section describe the solution of program (2) in the steady state (see Appendix for a description of the optimal path). These decision rules are interesting because the flow demand and stock demand perspectives have quite different implications. At the end of this section we expand the analysis to consider firms with a large market share, and we show how to modify (2) and the optimal decision rules.

3.1 Optimal marketing policy

In the steady state, the optimal level of messages M* meets the following first order condition for an interior optimum:

$$(3) \qquad (p - c_a - k_m \cdot m)\frac{\partial E}{\partial M} = (r + f*) \cdot c_m$$

where: p = price in period t. In many firms, such as mutual fund managers, the price is the product of a commission rate τ, and a base B which is the average base of the commission per customer. In that case, p ≡ τ·B.

∂E/∂M = sensitivity of entries to commercial messages that attract new clients.

Equation (3) takes into account the marginal benefit in the current instant of a message that attracts new clients, and the marginal cost of “renting” the investment in the message (which is c_m) for the current instant. The marginal benefit is simply the profit margin earned on members, which is price minus marginal administration cost minus marginal maintenance investment per member, times the number of new members attracted by the marginal message M.

Alternatively, if (r+ϕ*) is moved to the left hand side, equation (3) can be interpreted as taking into account the present value of the benefits earned by a message, with the investment in the message. This present value takes into account both the discount rate r and the rate of depreciation of the new clientele earned by message M, which is ϕ^*[9]. The optimal scale of marketing investment is $c_m \cdot M^*$.

Equation (3) implies, just as conventional models do, that a higher level of prices creates a direct economic incentive to invest more in marketing. The application of this principle may be tricky in some cases. Consider a case where a mutual fund manager experiences an increase in the funds under management, say because an increase in the prices of the securities held by the fund. Taking as given the commission rate τ, this should be interpreted as an increase in B, the average base of the commission per customer. The result is an increase in the firm's price p ($\equiv \tau \cdot B$), and equation (3) says that the manager has a higher incentive to invest in marketing.

In the steady state, maintenance messages m* meet the following first order condition for an interior optimum:

$$(4) \qquad (p - c_a - k_m \cdot m)\left(-\frac{\partial f}{\partial m} + \frac{1}{A}\frac{\partial E}{\partial m}\right) = (r + f^*) \cdot k_m$$

The left hand side of (4) shows the marginal benefit of increasing maintenance investments or quality of service, which is the marginal reduction in the rate of exits, plus the increase in the rate of entries, times the margin earned in the members attracted by the maintenance investment. The right hand side reports the marginal cost of "renting" the cost of investment in these messages (which is k_m) for the current instant. If both sides of equation (4) are multiplied by A, this equation can be interpreted on a dollar basis rather than on a dollar per member basis. Of course, there is also a present value interpretation when the discount rates are moved to the right hand side.

3.2 Optimal pricing strategy

In the steady state, the optimal price is set by:

---

[9] When the present value of the new clients obtained with the first message is below c_m, then the optimal marketing policy is zero investment (M*= 0). This may happen when messages are unproductive, when the operational margin ($p-c_a-k_m m$) is relatively low as compared to the cost of messages.

(5) $$\frac{p^* - c_a - k_m \cdot m^*}{p^*} = \frac{1 + (r / f^*)}{h^E + h^f}$$

where: ϕ^* = exit rate in the steady state (where dA/dt = 0).

$\eta^E \equiv -[\partial E/\partial p]\cdot(p/E)$ = *elasticity of entries to price*. Note that by definition $\eta^E > 0$ because $-\partial E/\partial p > 0$.

$\eta^\phi \equiv [\partial \phi/\partial p]\cdot(p/\phi)$ = *elasticity of exits to price*. In this case $\eta^\phi > 0$ as well.

Equation (5) is the form taken by the Lerner Rule in the stock demand model. This is visualized better by defining a "stock price elasticity" for a stock demand with two flow demand branches:

(6) $$h^{stock} \equiv (h^E + h^f) \cdot \left(\frac{f^*}{f^* + r} \right)$$

According to (5), the optimal margin is inversely proportional to the sum of the price elasticities of the two flow branches of the stock demand, $(\eta^E + \eta^\phi)$. For this reason, if one branch is price elastic and the other is not, overall demand is still price elastic. When at least one flow price elasticity is large, the firm looses the freedom to chose its price and is driven to choose a zero margin, becming effectively driven to perfect competition. In this case, the appropriate margin must discount the cost of maintenance marketing, in addition to administration marginal cost.

Equation (6) also shows that with a stock demand, the discount rate r has a direct positive influence over pricing policy. This is explained because at a higher discount rate r, the firm is relatively more interested in increasing current profits, even though this may bring a faster reduction in members through a higher exit rate and a lower entry rate.

The direct influence of the discount rate over prices is novel in microeconomic models, and therefore interesting. This has a host of implications that we do not develop in this paper. We just mention that an increase in the discount rate in the economy, say because of a tightening of monetary policy, leads to a direct increase in goods' prices[10].

The presence of the exit rate ϕ^* in (6) is explained by observing that $1/\phi^* \equiv AT^*$ is the average residence time of clients in the firm[11]. The average residence time can be interpreted as the aveage time during which clients are "captive" in this firm. This pushes towards the future the reduction in profits caused by a price increase, where they become discounted more heavily[12]. Thus, when the exit rate falls, and there is less "churning" ϕ, the clientele between firms, average residence time increases and the firm acts as if the price elasticity of demand had fallen, and raises prices. In the same way, if some regulation reduces churning to zero, the firm would act as if it had a fully captive clientele and, provided the discount rate is positive, would raise prices indefinitely even if the sum of price elasticities is large.

In a number of industries that face a stock demand, firms charge commission *rates* such as a percentage of assets under administration, rather than prices measured in dollars. For example, in mutual fund managers the price is the product of a commission rate τ and a base B, so that $p \equiv \tau \cdot B$. In this case, equation (5) also asserts that when an exogenous trend or shock raises the base B, the optimal commission rate τ^* falls inversely proportionately. Moreover, the fall in the commision rate should be an instantaneous jump, even if the pricing policy has not achieved the steady state, because an economic model like (2) fixes an optimal path for the product $\tau(t) \cdot B(t)$, without specifying how it is distributed between τ and B. This prediction illuminates the nature of competition in the equity mutual fund managers in the U.S., because those firms did not reduce commission rates when the stock market boom of the 1990's tripled the assets under management.

In the conventional models of marketing investment, such as (1), the number of messages does not influence the optimal level of price, except when it influences the price elasticity of demand. This is a problem for flow demand models, because expenditure in messages effectively allow the firm to choose the nature of its service, because an increase

---

[10] In any Keynesian or new Keynesian macroeconomic model, this would create a new direct channel through which monetary policy would constrain aggregate supply in the short run.

[11] In any steady state the entries and exits are equal: $E = \phi A$.

[12] To go through, this argument needs the discount rate to be positive, and this is precisely what we find in (6), because when $r = 0$, the exit rate does not influence the effective price elasticity of demand.

in prices can be compensated by an increase in messages. But if this is the case, then the number of messages delivered affect the firm's margin and may influence it.

In the stock demand model analyzed here, the number of messages that attract new clients do not influence pricing directly. However, expenditure in maintenance messages do influence prices directly. As shown by (5), their cost is part of the regular services provided to clients and reduce the margin directly. This is a major difference between m and M. Maintenance messages allow the firm to choose the nature of the service provided to members, because an increase in prices can be compensated by an increase in maintenance messages, and in the stock demand model, their cost do show up in the optimal margin.

Instead, expenditure in messages that attract new clients are not a "necessary cost" of doing business, but rather are an investment in expanding the clientele whose magnitude is chosen by the firm according to the prospects of that investment. Just as the fixed cost F does not appear in (5), investment $c_m \cdot M$ does not appear either.

The implication is that an increase in marketing investment in attractive messages cannot *cause* an increase in prices. Conversely, a drop in this sort of investment should not cause a reduction in prices. Only an increase in maintenance marketing expenditures does increase effective marginal costs and is expected to increase prices. The reason is that the maintenance investment cost, is in effect an addition to marginal costs of production. This result contradicts assertions related to the new private pension fund management industries of Argentina and Chile, which say that when these firms increased marketing investment in the mid 1990's, commissions were forced to increase[13]. This is wrong, because most of the marketing investment made by those firms was of the attractive type, whereby salespeople were paid only for attacting new members.

3.3 Combined first order conditions

Combining (3) and (4) with (5) it is easy to show that the steady state optimal policy meets the following conditions, when it is an interior solution:

[13] For example, the Chilean Superintendent of AFP for 1982-90 proposed to introduce regulations "to reduce churning, and reduce marketing investments, allowing in that way a fall in commissions" (Ariztía, 1997).

(7a) $$\frac{c_m \cdot M^*}{p^* \cdot A} = \frac{e_{E,M}}{h^E + h^f}$$; (7b) $$\frac{k_m \cdot m^* \cdot A}{p^* \cdot A} = \frac{e_{E,m} + e_{f,m}}{h^E + h^f}$$

where $\varepsilon_{E,m} = (\partial E/\partial m)(m/E) > 0$ is the elasticity of entries to maintenance expenditures, $\varepsilon_{\phi,m} = -(\partial\phi/\partial m)(m/E) > 0$ is the elasticity of the exit rate to maintenance expenditures, and $\varepsilon_{E,m} = (\partial E/\partial M)(M/E) > 0$ is the elasticities of entries to atttractive messages.

The left hand side of (7) is the optimal ratio of each type of marketing expenditures to sales, in the steady state. The right hand side shows in the numerator, the elasticities of entries and exits to messages, and in the numerator, the price elasticities. Equation (7) is analogous to the optimal marketing sales ratio for a conventional model like (1) (see Berndt, 1991 for a summary).

The stock demand model allows a novel interpretation of exit rates. Combining (3) and (5) to eliminate p*, we may obtain the following “optimal churning rate”:

(8) $$f^* = \frac{r}{(h^E + h^f) - 1} + \left(\frac{c_a + k_m \cdot m^*}{c_m} \right) \cdot \frac{\partial E / \partial M}{(h^E + h^f) - 1}$$

This equation states that it is optimal to increase the churning rate at the firm enough to make it unprofitable to raise further the investment in marketing that attracts entry. The average residence time of new members falls to the level where it is to short to justify further increases in the marketing investment rate in attractive messages. In equation (8), the firm takes into account that as churning increases, the effective price elasticity increases and it is desirable to reduce price, which in turn reduces the benefit of investing in marketing messages and reduces further the optimal churning rate.

3.4 Firms with a large market share

For a firm with a substantial market share, model (2) must be extended because the proximity of the ceiling placed on the firm’s clientele by the universe of potential clients reduces the effectiveness of a given level of marketing messages, in terms of attracting new clients. Thus, the flow of entries is smaller for a given marketing investment. Following the literature[14], we take into account the impact of market share in total entries and replace the first restriction in problem (2) by:

(9a) $$\frac{dA}{dt} = C(p,M,m)\cdot\left(1-\frac{A}{A^U}\right) - f(p,m)\cdot A$$

(9b) $$\equiv C(p,M,m) - \left(f(p,m) + \frac{C(p,M,m)}{A^U}\right)\cdot A$$

where C(p,M,m) = flow of contacts with potencial entrants.

The change in (9a) is that entries depend negatively on the current market share of the firm. The idea is that any given marketing and pricing effort generates C contacts with clients that can potentially enter, but a proportion of those contacts fail because the client already has a contract with the firm. Equation (9a), following the literature[15], postulates that this proportion is equal to the market share of the firm. The market share term is effectively constant for firms with a small market share (it remains near zero), but not for large firms. Equation (9b) reinterprets this change as an increment to the exit rate. It says that when the clientele increases, exits for a given depreciation rate increase as before, but in addition the number of contacts that fail for a given marketing investment rate also increases.

Another aspect of a large market share is that the universe of potential members, A^U, may change in response to the actions of the firm. For example, a lower average price and higher total marketing investment may increase the size of the market. This effect can be specified in the stock demand spirit, i.e. proposing a “slow” change in the universe of potential clients, or in the flow demand spirit, i.e. assuming that changes in the controls affects the universe of potential clients instantaneously. To show the versatility of our approach, we adopt the flow approach and propose the model:

[14] See Nerlove and Arrow (1962), the model by Gould (1970) and the summary by Sethi (1977), all cited by Tu (1984).

[15] Gould (1970) analyzed the case where entries are proportional to the product $A\cdot(A^U-A)$, which may be justified when the contacts are made mainly by recommendation of current clients, not by marketing messages paid by the firm. Gould only solved the simplified case where price p and maintenance messages m (or quality of service) are exogenous, which is quite restrictive. See a summary in Tu, 1984, p. 334-7.

(10) $$A^U = A^U(\bar{p}, \sum M_j, \bar{m}) = A^U(ap + (1-a)\bar{p}_{j\neq i};\ M + \sum_{j\neq i} M_j;\ am + (1-a)\bar{m}_{j\neq i})$$

where $\alpha = A/A^U$ is the market share of the firm and $\partial A^U/\partial p < 0$, $\partial A^U/\partial M > 0$, $\partial A^U/\partial m > 0$. In the symmetric case where all firms in the industry choose the same policies, it is also the case that $\partial A^U/\partial A = 0$, a condition we assume aplicable.

Working through the first order conditions for the steady state, the combined consequences of (9) and (10) are as follows:

(i) the effective depreciation rate of the clientele rises from ϕ to $(\phi + C(p,M,m)/A^U)$. In the steady state, this increase in the effective exit rate simplifies to an increase from ϕ to $\phi/(1-\alpha)$, where $\alpha = A/A^U$ is the market share of the firm.

(ii) Regarding the optimal pricing policy, equation (5) must be replaced by:

(5’) $$\frac{p^* - c_a - k_m \cdot m^*}{p^*} = \frac{1}{\left((1-a)\left[h^E + h^f\right] + a^2 h^{AU,p}\right) \cdot \left(\frac{f}{r(1-a)+f}\right)}$$

where $\eta^{AU,p} \equiv -[\partial A^U/\partial p] \cdot (p/A^U)$ = elasticity of potential market size to average price.

Equation (5’) says that as the market share of the firm approaches 100%, the profit maximizing margin is given by the standard Lerner Rule, i..e. is inversely proportional to elasticity of potential market size to average price. The influence of the churning rate and the flow elasticities of individual demand fall to zero.

In the case of a mandatory market, such as pension fund management services in the mandatory social security systems of Argentina, Chile[16] and Mexico, $\eta^{AU,p} = 0$. In this case, equation (5’) says that as the market share of the firm increases, the optimal policy becomes an infinite price, which is easily understood because the clientele becames fully captive. Thus, an increase in industry concentration can be much more damaging to consumers in mandatory markets than in voluntary ones.

(iii) Regarding optimal marketing policy, (3) changes to:

---

[16] For a detailed description of the Chilean case, see Diamond and Valdés-Prieto (1994).

$$(3') \qquad (p - c_a - k_m \cdot m)\left[(1-a)\cdot\frac{\partial E}{\partial M} \;+\; a^2 \cdot f \cdot \frac{\partial A^U}{\partial \Sigma M}\right] = (r(1-a)+f)\cdot c_m$$

where [∂A^U/∂ΣM] = sensitivity of potential market size to total marketing effort.

Equation (3') shows that as the firm raises its market share, the response that becomes more important is that of potential demand, and the churning rate loses its influence as it cancells out. In the case of large firms that serve a mandatory market, where [∂A^U/∂ΣM] = 0, equation (3') says that as the market share of the firm increases, the marginal benefit of marketing falls to zero while the marginal cost remains at c_m, so optimal marketing policy becomes zero investment (M* =0). Thus, we should expect an increase in market concentration to reduce the expenditure on salespeople by each one of the firms involved, even in the absence of collusion.

(iv) Regarding optimal service policy, or marketing maintenance, (4) changes to:

$$(4') \qquad (p - c_a - k_m \cdot m)\left((1-a)\left[-\frac{\partial f}{\partial m} + \frac{1}{A}\frac{\partial E}{\partial m}\right] + a^3 f \frac{\partial A^U}{\partial \overline{m}}\right) = (r(1-a)+f)\cdot k_m$$

where [∂A^U/∂m] = sensitivity of potential market size to average quality of service.

Again, equation (4') shows that as the firm raises its market share, the potential demand response becomes critical and the churning rate cancells out.

In the case of a mandatory market, where [∂A^U/∂m] = 0, equation (4') says that as the market share of the firm increases, the optimal quality of service or maintenance becomes zero (m* =0). This is because the marginal benefit of quality or maintenance falls to zero, while the marginal cost remains at k_m, so optimal quality policy is m* =0. Thus, we should expect an increase in market concentration in the mandatory social security systems of Argentina, Chile and Mexico, to bring with it a fall in the investment in the main dimension of service quality, which is financial performance, i.e. the expected rate of return of the funds under management, for a given level of risk. In other words, we should expect an increase in market concentration to induce the attrition of the investment management teams hired by these firms, and to substitute the best talent for normal and cheaper talent. The reason is that the marginal benefit of talent goes to zero as the firm's share grows.

4. The Nash equilibrium in the steady state

The game the firm faces is an oligopoly where N firms choose sequences of actions over time. We ask if there exists a perfect equilibrium where the actions of each firm are constant over time, i.e. if there is any Nash equilibrium that is a steady state. In a candidate perfect equilibrium, the actions prescribed by the steady state must be part of a Nash equilibrium between the firms in the continuation subgame where they choose actions simultaneously. If this were not the case, program (2) would be an inadequate representation because rivals would choose to change their actions over time, and the entry and exit functions in (2) should include explicitly the time argument. This would imply that the steady state rules identified in section 3 are not optimal. For the assumed constancy of rivals' actions to be credible, we have to check that the stratgies prescribed for the subgame that begins at the steady state are themselves a Nash equilibrium.

This section investigates these questions for an industry with N identical firms, i.e. which face the same stock demand and cost conditions. In this setting, a natural candidate equilibrium is the symmetric Nash equilibrium where all firms take the same actions[17].

It is possible that multiple asymmetric Nash equilibria also exist in the continuation subgame. For example, in the standard Cournot one-shot game, there exist multiple asymmetric Nash equilibria when the best response curve has a slope smaller than 45º in absolute value[18]. In that case, the points where one firm stops production are also Nash equilibria, though asymmetric.

In the presence of multiple Nash equilibria, game theory usually compares them using some refinement. In the Cournot one-shot game, the condition for a multiplicitly of equilibria is the same condition for the hypothetical tatonnement process of adjustment around the equilibrium to be "unstable", in the sense of Cournot (1838). This fact suggests the following refinement: if the symmetric Nash equilibrium is unstable under the tatonnement process that follows a movement out of equilibrium by a single firm, the

[17] Of course, there may exist other perfect equilibria in this game, some dynamic and some asymmetric.

[18] This is impossible for linear market demand, but can happen for other shapes of market demand.

candidate equilibrium should be rejected. We are not suggesting in any way that the tatonnement process occurs in calendar time. We simply subject the candidate equilibrium to a refinement: if it not stable, then we declare it is unconvincing or unreasonable.

4.1 The slope of the best response curve

An important question regarding investment in M messages is whether a firm should increase or decrease its investment when a rival invests more, i.e. whether those marketing investments that increase entries are strategic complements or strategic substitutes[19]. In the case of a stock demand, one is tempted to answer that when rivals invest more in marketing and raise our firm's exit rate, the firm should respond by *increasing* its investment as well, in order to "replace the lost clients". On the other hand, if rival actions raise the firm's exit rate, then the depreciation of the new clientele proceeds faster and it is less valuable, so the firm might respond by *reducing* marketing investment.

To simplify the analysis, this section assumes that post-sale marketing (m) - or quality of service - is held at constant levels by all firms. In order to focus on the marketing actions that attract entries, i.e. on M, we also assume that prices are fixed. The fixed price case applies to industries under semicollusion, i.e. industries that have colluded in prices but compete in marketing investment. The empirical importance of this case has been defended by Schmalensee (1976). In addition, the fixed-price assumption is relevant for industries where the authorities regulate prices. Finally, fixed prices are the outcome of models where firms self-regulate prices in order to reduce the risk of a policy intervention that would reform the market rules and reduce their profits (Valdés-Prieto, 2001).

In program (2), expanded to the case of a firm with market share $\alpha > 0$, the actions of the N-1 rivals are included in the functions $E_i = (1-a)\cdot C_i(M_i, p_i, m_i, \overline{m}_{i\neq j}, p_j)$, $f_i = f_i(m_i, p_i, \sum_{j=1}^{N-1} M_j, p_j)$ and in equation (10) regarding A^U.

Let us consider the case where out of the N-1 rivals in the market, only rival j increases its pre-sale marketing investment. This allows us to rewrite (3') as:

$$(11) \qquad (p - c_a - k_m m^*)\left[(1-a)^2 \frac{\partial C}{\partial M}(M_i^*) + \ a^2 f(M_j)\frac{\partial A^U}{\partial \Sigma M}\right] = r(1-a)c_m + f(M_j)c_m$$

Total differentiation of the original version of (11) from the Hamiltonian with respect to M_i and M_j shows that marketing investments are strategic substitutes, because the right hand side of (12) is negative:

$$(12) \qquad \frac{dMi^*}{dMj} = \frac{-f'}{(-H_{MM})} \cdot \frac{\partial C}{\partial M}(p - c_a - k_m m)\left(\frac{1-a}{r(1-a)+f}\right)^2 + a^2 f \frac{\Omega}{(-H_{MM})} < 0$$

The first term in this expression is *negative* because $H_{MiMi} < 0$ by the S.O.C. (H is the Hamiltonian in the Appendix, expanded for the case of a large firm) and $\phi'_{ij} = \partial\phi i/\partial Mj > 0$ by the assumption that a rival's marketing increases the exit rate. The second term is generally negative[20], as well, except at points where the universe of potential clients is very convex to industry marketing. In industries with mandatory demand this is impossible.

The economic implication is that in most cases, when rival marketing raises the firm's exit rate, the firm responds optimally by *reducing* its own marketing investment.

4.2 Multiple asymmetric Nash equilibria in marketing messages

The size of the marketing response identified in (12) is critical for the existence of multiple asymmetric Nash equilibria in marketing. Consider an asymmetric point in the extension of figure (2b) for three or more firms (N=3). Given that firms 1 stops marketing, it is optimal for firms 2 and 3 to invest in marketing to draw the rival's clients to themselves over time. And given that these firms' strategies keep the exit rate at firm 1 high enough, it is optimal for firm 1 not to invest in marketing. In the case N = 2 the assymmetric equilibia would be more complicated, because after the single rival actually

---

[19] See Tirole (1988), p. XXX

[20] $\Omega \equiv \frac{p - c_a - k_m m}{r(1-a)+f}\left[\frac{-2}{A^U}\left(\frac{\partial A^U}{\partial \Sigma M}\right)^2 - \frac{f'}{r(1-a)+f} \cdot \frac{\partial A^U}{\partial \Sigma M} + \frac{\partial^2 A^U}{\partial \Sigma M^2}\right]$

disappears, there is a discontinuity in the dominant firm's policy over time: once it has a monopoly, it is best to reduce marketing investment to zero. By concentrating in cases with $N \geq 3$, we avoid that problem.

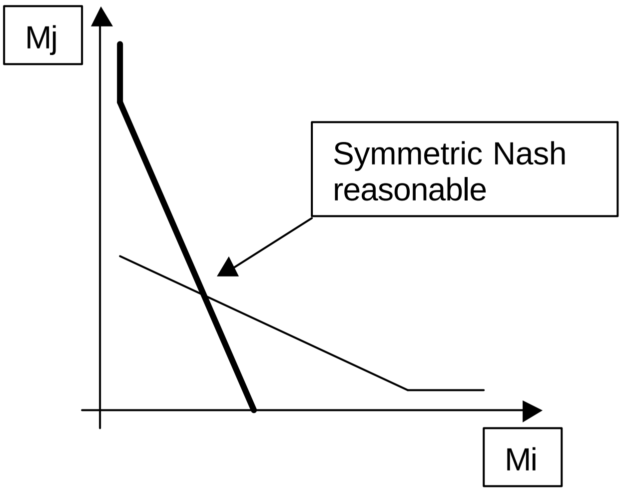

Figure 2a: Symmetric Nash is reasonable.

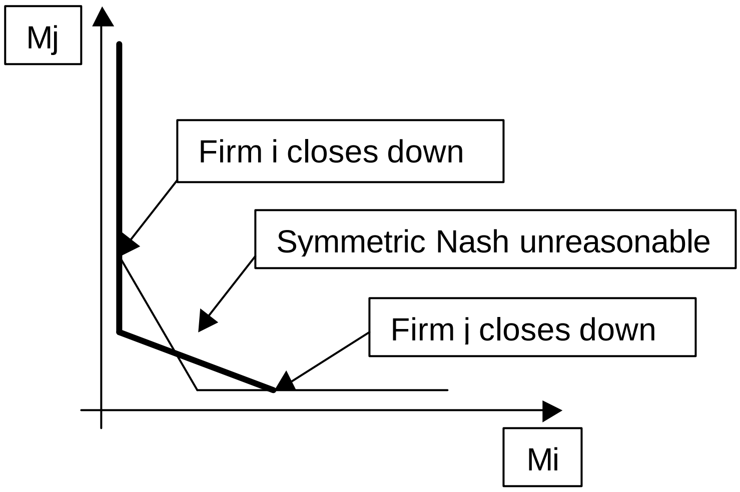

Figure 2b: Asymmetric and multiple Nash eq.

To choose between equilibria in figure 2b, we adopt as a refinement the criterion that candidate Nash equilibria must be stable in Cournot's tatonnement process. It is easy to see that the symmetric equilibrium fails this test and the asymmetric ones pass it in figure 2b: Consider a case where the response is large, i.e. where an increase in a rival's marketing effort increases the exit rate of our firm, and our best response is to reduce our own marketing effort considerably enough. If this happens within a tatonnement process, the rival will see its own exit rate drop by enough to induce it to increase its marketing effort further. This process would cumulate until our firm's best response is be to stop marketing and to allow exit to erode its market share until closedown. A candidate equilibrium with this outcome is not deemed reasonable.

To apply this refinement, we propose the following tatonnement process that starts in the symmetric Nash equilibrium and evolves in discrete time as follows:

t=1: firm 1 raises its marketing investment by dM_1;

t=2: the N-1 rivals play a new symmetric Nash equilibrium between them;

t=3: firm 1 responds to the changes of its rivals, and so on.

As is well known, this process is stable if the absolute value of $dM1_{(t=3)}/dM1_{(t=1)}$ is smaller than 1.

When the symmetric Nash equilibrium does not pass this refinement, we are left with the asymmetric ones. The multiplicity of those equilibria is a problem because there is no clear reason for the firms to focus on one of them. Moreover, the identity of the chosen equilibrium has significant profit implications for the firms. Therefore, there is no reason to preclude outcomes that are not Nash equilibria, at least for some time. For example, one can imagine a war of attrittion in which several firms invest strongly in marketing in order to maintain alive a chance of pushing the industry to an equilibrium where they make profits During the war of attrittion several firm can make losses for a random period of time. This may describe the "marketing wars" observed in the Argentinean and Chilean pension fund management markets until 1997.

Although the outcome of these games is the close down of at least one firm, this does not means that some firms must necesarily stop marketing and allow their clientele to exit over time. An alternative route to close down is a merger or acquisition at a point in time much before close down.

Mergers can be profitable for several standard reasons. One applies when the merger allows better exploitation of economies of scale or scope. In the model of this paper, the marginal production cost has been assumed constant, and the presence of a fixed cost F has been postulated. This implies economies of scale that make a merger gainful. Of course, this does not imply that the industry is a natural monopoly, because the model also allows for product differentiation (η^E and η^ϕ are not ∞). Another standard reason for mergers is that in the absence of collusion, they reduce competition both in prices and marketing, allowing higher profits for the merged firm (and also for the other firms).

The game analyzed here introduced a novel reason for mergers: due to the economies of scale postulated in the cost function, it is always more profitable to sell the closing firm as early as possible than holding out until all clients are lost. As the size of the clientele dwindles, the average cost of production rises without bound because the fixed cost must still be paid.

This argument suggests that one of the asymmetric Nash equilibria shown in figure (2b) might be reached as soon as the nature of the situation is recognized by those firms that will choose to close down, and this will happen through a merger wave induced by this novel reason[21]. Thus, the "war of attrition" may end abruptly with a merger wave. This may explain the merger wave observed in the Argentinean and Chilean pension fund management markets during 1996-98.

An important issue is whether the likelihood of such consolidation increases when concentration falls or when concentration rises. In the former case, an atomistic industry structure would be inlikely. In the second case, the industry structure must be either atomistic or a complete monopoly, because the intermediate cases would push towards consolidation. The complete monopoly outcome also requires perfectly symmetric demand and cost conditions, because asymmetry may allow some firms to remain as niche players and only a dominant firm structure will obtain.

4.3 The case of mandatory demand

Application of this tatonnement process to the case of a general N-firm symmetric industry is complicated. Therefore, we study an important special case: the one where demand is mandatory and A^U is not influenced by marketing effort. This may represent the case of mandatory education coupled with school vouchers, and the case of "privatized" social security where workers are mandated to choose among registered fund management firms, as in Latin America.

We determine the stability of the tatonnement process for this case. When A^U is not influenced by marketing effort, and starting directly from (11) we find that:

$$(13) \qquad \frac{dM1(t=3)}{dMj \neq 1(t=2)} = \frac{-f' \cdot c_m}{(-C_{MM})(p - c_a - k_m m^*)(1-a)^2} < 0$$

Equation (13) shows that what matters for firm's 1 exit rate ϕ is the absolute increase in rivals' messages Mj, not the increase relative to the total number of rival

[21] When cost and demand are symmetric, there is not enough information to identify the firm that will close up. In real-life situations asymmetries always exist, and they allow identification of the firms that will close: those with the weakest position, either because of higher costs or weaker demand niches.

messages (Mj + ΣM_k). Provided ϕ' remains constant, the right hand side of (13) does not fall towards zero as the number of rivals increases.

We also ask what happens in t=2 of the tatonnement process: when firm 1 increases its own marketing investment by a small amount, the N-1 symmetric rivals respond by resetting their marketing investments simultaneously, according to a new Nash equilibrium between them. Recalling that the marketing investment of a typical rival does not enter its own exit rate, and that N ≥ 3, the symmetric Nash equilibrium in messages M is represented by (3'), which simplified by the fact that A^U is fixed, amounts to:

$$(14) \qquad (p - c_a - k_m \cdot m)\left[(1-a)\cdot\frac{\partial E}{\partial M}(M^N) \; + 0\right] = r(1-a)\cdot c_m + f(M1 + (N-2)M^N)\cdot c_m$$

where M^N = amount of marketing investment in the symmetric Nash equilibrium between N-1 rivals, during state t=2 of the tatonnement process.

Differentiating (14) we find that each of the rivals responds to an increase in M1 by reducing its marketing investment at the rate:

$$(15) \qquad \frac{dM^N(t=2)}{dM1(t=1)} = \frac{-1}{(N-2) + (-C_{MM})\dfrac{(p - c_a - k_m m^*)}{c_m \cdot f'}(1-a)^2} < 0$$

where the identity $E_{MM} \equiv (1-\alpha)C_{MM}$ was considered.

Equation (15) shows that as the number of rivals increases, the impact of a change in firm 1's messages on each rivals actions falls towards zero. Now we check the stability of this tatonnement process. Note that:

$$(16) \qquad \frac{dM_1(t=3)}{dM_1(t=1)} = \frac{d\Sigma M^N(t=2)}{dM_1(t=1)}\cdot\frac{dM_1(t=3)}{d\Sigma M^N(t=2)} = (N-1)^2\,\frac{dM^N(t=2)}{dM_1(t=1)}\cdot\frac{dM_1(t=3)}{dM^N(t=2)}$$

Let us explain the $(N-1)^2$ term in (16): in t=2 all N-1 rivals of firm 1 reduce their marketing investment by the same amount, so the total response is (N-1) times the individual response reported in equation (15). Then, in t =3, the best response of firm 1 is to increase its own marketing investment by (N-1) times the response to a single rival's reduction in messages, which is reported by equation (13).

Introducing (15) and (13) into (16), and using that $\alpha = 1/N$, we find that:

$$(17) \qquad \frac{dM1(t=3)}{dM1(t=1)} = \frac{N^2}{K \cdot \left[N - 2 + K\left(\frac{N-1}{N}\right)^2 \right]} \qquad \text{where } K \equiv \frac{(-C_{MM})}{f'} \cdot \frac{(p - c_a - k_m m^*)}{c_m}$$

As is well known, this tatonnement process is stable if the right hand side of (17) is smaller than 1, and unstable in the reverse case. Analysis of equation (17) shows the following results[22]:

(a) When $K < (-1+\sqrt{17})(8/9) \approx 3{,}5134937$ the tatonnement process is unstable for any level of concentration. This implies that the symmetric Nash equilibrium in messages is discarded by the proposed refinement, and a merger wave raises concentration until a monopoly obtains (for symmetric demand and cost conditions). Thus, the optimal policies proposed in section 3 are not a perfect equilibrium.

(b) For K > 3,5134937, the symmetric atomistic market structures are unstable. For example, for K=5, any number of firms larger than 5 generates an unstable tatonnement process. This implies that the symmetric Nash equilibrium in messages is discarded by the proposed refinement, and a merger wave raises concentration until the region described in (c) is reached.

(c) For K > 3,5134937, the concentrated symmetric market structures are stable[23]. As the symmetric Nash equilibrium in messages meets the proposed refinement, the proposed steady state is a Nash equilibrium and the optimal policies proposed in section 3 are a perfect equilibrium.

(d) As K grows without bound, any non-atomistic market structure becomes stable. For example, for K=20, the symmetric market structures with $N \leq 30$ are stable, allowing for substantial deconcentration of the industry structure.

---

[22] Note that equation (17) is valid for $N \geq 3$ only, as we assumed this in its derivation.

Summing up, we have found that once the market becomes concentrated enough, the incentive for a merger wave discussed section 4.2 disappears. However, when the value of parameter K is small enough, mergers must proceed until monopoly is achieved (provided no niche players exist).

The economics of stock demand industries with mandatory demand, analyzed in this section, are governed by the value of parameter K. This in turn is governed by various factors, of which the most interesting seem to be:

(i) When $\phi' \approx 0$, a condition that may obtain more easily when a firm believes its service to be differentiated enough to make its exit rate relatively insensitive to the marketing investments of rivals, the value of K is high and a relatively dispersed market structure is stable and part of a perfect equilibrium. This suggests a link between the degree of product differentiation in the industry and its concentration: once the degree of differentiation becomes high enough, K may be driven above its critical value and an dispersed industry structure might become stable.

However, this may not be socially desirable. Product differentiation also brings with it a reduced sensitivity of the exit rate to prices. The impact of such a change is shown in equation (5) to lead to higher prices in the steady state, so consumers are hurt. Because of this loss to consumers, regulatory encouragement to increase product differentiation would be specially inappropriate public policy in a mandatory market[24].

(ii) When $|C_{MM}| >> 0$, a condition that is likely to obtain when the marginal productivity of messages drops strongly at a threshold level, the value of K is high. This suggests that a collusive agreement that imposes quotas on the number of messages issued by each firm, such as quotas on salespeople, or a regulation that imposes such quotas, makes a relatively dispersed market structure stable and part of a perfect equilibrium.

Again, this may not be socially desirable. Message quotas can also reduce the price elasticity of flow demand, as shown by the empirical study of Marinovic (2000) for the Chilean pension fund management mandatory market. That study shows that the sensitivity

---

[23] Specifically, those N small enough so that $N^4 - KN^3 + (2K-K^2)N^2 + 2K^2N - K^2 < 0$.

[24] It seems impossible to engineer regulations that reduce the sensitivity of the exit rate to rivals'marketing investments and at the same time increase the analogous sensitivity to prices.

of exits to price differences is proportional to the number of salespeople (messages), provided the firm charges a smaller price than competitors. As shown in equation (5), the impact of such quotas is higher prices in the steady state, so consumers are hurt. Because of this loss to consumers, regulatory encouragement to collusion in messages would be specially inappropriate public policy in a mandatory market.

5. Final comments

This paper presents a model of competition in “stock demand” markets, which seems appropriate for subscription markets, the financial services industries, cell phones, mandatory education financed with vouchers and manadatory “privatized” social security. In these markets the clientele with contracts evolves slowly. It is shown that in a perfect equilibrium, the unique optimal price path leads to a steady state where the effective price elasticity depends on the rate of depreciation of the clientele through exit, on the discount rate, and on the elasticity of both entry and exit flows to prices. It is found that a “large” firm that faces a stock demand in a mandatory market, such as those for pension management services, the optimal margin rises indefinitely as market share goes to 1, and the optimal fund management effort goes to zero.

A perfect symmetric equilibrium does not exists always, however. Under some parameter conditions there can be multiple asymmetric Nash equilibrium in marketing investment, which are more reasonable than the symmetric equilibrium because the latter does not meet a refinement we propose. The asymmetric equilibria lead to a war of attrition and to mergers that concentrate the industry. This model may explain the marketing wars and the merger waves observed in the mandatory pension fund management markets for “privatized” social security services of Argentina and Chile in the mid 1990’s.

In the case of mandatory demand, it is shown that for an important region of parameter values, only concentrated industry structures allow a perfect symmetric equilibrium to exist, while excessively atomistic structures encourage mergers. This result is analogous to the Sutton (1991) result, where a more concentrated industry is more likely to achieve a viable equilibrium, than a relatively more atomistic industry. The empirical backing summoned by Sutton (1991) suggests that this model may be of wider interest.

In mandatory markets, regulatory intervention or collusive agreements may seek changes in the critical parameter values to achieve a viable perfect equilibrium for more dispersed industry structures. It is argued that these interventions - that artificially encourage product differentiation or establish quotas on marketing messages - also encourage higher prices that hurt consumers, and thus could be inappropriate policy in a mandatory market. This is the case because an alternative organization of industry exists for mandatory markets that may offer a more desirable social outcome: the government may call for a bidding competition to supply the mandated services to designated sets of consumers, favoring the firms that offer the lowest prices. This approach, followed by Bolivia and Panama for the pension fund management industry in 1997 and 2000, has achieved substantially lower prices for consumers.

References

Ariztía, J. (1997) "Análisis sobre los traspasos en el sistema previsional", en J. Ariztía, editor, *AFP: Las Tres Letras que Revolucionan a América*, CIEDESS, Santiago, Chile, p. 311-332.

Berndt. E. (1991) "Causality and Simultaneity between Advertising and Sales", capítulo 8 en *The Practice of Econometrics: Classic and Contemporary*, p. 361-448; Addison-Wesley Publishing Co., Reading, Massachusetts.

Diamond, P.A. and S. Valdés-Prieto (1994) “Social security”, Chapter 6 in Bosworth, B., R. Dornbusch and R. Labán (eds.) *The Chilean Economy: Policy Lessons and Challenges*, The Brookings Institution, Washington D.C.

Gould, J.P. (1970) "Diffussion processes and optimal advertising policy", en E.S. Phelps, editor, *Microconomic Foundations of Employment and Inflation Theory*, Norton, New York.

Marinovic, I. (2000) “Estimación de la Demanda por AFP en Chile: 1992-1997”, Thesis in Economics, Institute of Economics, Catholic University of Chile, Santiago.

Nerlove, M. and K. Arrow (1962) “Optimal advertising policy under dynamic conditions”, *Economica* 39, p. 129-142, May.

Rizzo, J. (1999) "Advertising and competition in the Ethical Pharmaceutical Industry: the case of Antihypertensive drugs", *Journal of Law and Economics* 42 (1), Part 1, p. 89-116.

Roberts, M. and L. Samuelson (1988) "An empirical analysis of dynamic, nonprice competition in an oligopolistic industry", *Rand Journal of Economics* 19, No. 2, Summer, p. 200-220.

Schmalensee, R. (1976) "A model of promotional competition in oligopoly", *Review of Economic Studies*, vol. 43, No. 4 (October), p. 493-507.

Schmalensee, R. (1977) "Comparative static properties of regulated airline oligopolies", *Bell Journal of Economics*, v. 8, Autumn, p. 565-576.

Sethi, S.P. (1977) "Dynamic optimal control models in advertising: a survey", *SIAM Review*, 19 (4), p. 685-725, October.

Sutton, J. (1991) *Sunk Costs and Market Structure: Price competition, Advertising and the Evolution of Concentration*, MIT Press, Cambridge.

XXX Shipping industry

Tirole, J. (1988) *The Theory of Industrial Organization*, The MIT Press.

Tu, Pierre (1984) *Introductory Optimization Dynamics*, Springer-Verlag, New York.

Valdés-Prieto, S. (2001) "Los mercados de AFP de Argentina y Chile: Semicompetencia", mimeo, Instituto de Economía, P. Catholic University of Chile, Santiago, Chile, March.

Appendix: Solution to the optimization problem (2)

Assuming the second, third and fourth restrictions in program (2) are slack, the Hamiltonian is:

(A.1) $H(p, M, m, A, \lambda) = e^{-rt}\cdot\{ (p-c_a)\cdot A - F - c_m\cdot M - k_m\cdot m\cdot A + \lambda\cdot[E(p,M,m) - \phi(p,m)\cdot A] \}$

$\lambda(t)$ is interpreted as the value of a new client attracted or lost at date t. This is the effect of a unit change in the clientele in the present value of the firm as of time t. As of time 0, its value is $\lambda\cdot e^{-rt}$.

First Order Conditions

From the Maximum Principle:

(A.2 a) $\partial H/\partial A = -d(\lambda\cdot e^{-rt})/dt$ => $(p-c_a- k_m\cdot m)/\lambda - \phi + d\ln\lambda/dt = r$

(A.2 b) $\partial H/\partial p = 0$ => $A = \lambda\cdot(- \partial E/\partial p) + \lambda\cdot (\partial\phi/\partial p)\cdot A$

(A.2 c) $\partial H/\partial M = 0$ => $c_m = \lambda \cdot \partial E/\partial M$

(A.2 d) $\partial H/\partial m = 0$ => $k_m \cdot A = \lambda \cdot [(\partial E/\partial m) + (^{-}\partial\phi/\partial m)\cdot A]$

(A.2 e) $\partial H/\partial(\lambda \cdot e^{-rt}) = dA/dt$ => $dA/dt = E(p,M,m) - \phi(p,m)\cdot A$

The left hand side of (A.2 a) shows the total return over capital invested in clients as of time t, which should be equal to alternative return r along the optimal path. This total return is the sum of the dividend, i.e the operating margin left by a client after substracting the costs of service quality or maintenance marketing, expressed as a proportion of its value as a percentage of its value λ, minus "depreciation" through exits, plus the "capital gain or loss" whose sign and size depends on whether the value λ of clients is increasing or falling over time. Using (A.2a) to determine λ we find the following value of a client at time t in the optimal path:

(A.2 a') $\lambda = (p - c_a - k_m \cdot m + d\lambda/dt)/[r + \phi]$

This equation says that the value of a client should consider dividends and capital gains and that the appropriate discount rate is the sum of the alternqative return and the exit rate. If sme regulation reduces the exit rate ϕ, the value of each client increases instantaneously.

Equation (A.2b) recommends an increase in price p until the gain from a higher income from the current clientele is fully outweighted by the sum of two losses: the reduction in entries and the increase in exits. Equation (A.2c) recommends an increase in marketing messages that attract new clients as long as the the value of the new clientes attracted is above the marginal cost of the message.

Equation (A.2d) recommends that the quality of service, or the level of maintenance marketing messages, be raised as along as the increase in marginal cost of production remains below the marginal benefit, which in turn is the sum of value of the clients attracted and the value of the clients that are not lost.

Of course, these first order conditions are valid for interior optima only. If, for example, $A \cdot k_m >$ $\lambda \cdot [(\partial E/\partial m) + (^{-}\partial\phi/\partial m)\cdot A]$, then m* = 0, i.e the firm shouls stick to standard quality of service and avoid maintenance marketing.

Border Conditions

In program (2), the border conditions are the initial condition $A(0) = A_o$ and the terminal or transversality condition, which is (A.3.a). Equation (A.3b) arises from (A.3a) and the second restriction A(t) = 0.

(A.3.a) $\lim_{t\to\infty} \lambda^*(t)\cdot e^{-rt}\cdot A^*(t) = 0.$

(A.3.b) $\lim_{t\to\infty} \lambda^*(t)\cdot e^{-rt} \geq 0$ (see Tu, 1984, p. 134).

Condition (A.3.a) is met if and only if r >0 and the product $\lambda^*(\infty)\cdot A^*(\infty)$ is bounded. Condition (A.3.b) is met only if r >0 and $\lambda^*(\infty)$ is bounded. Therefore, the border conditions require that the discount rate is positive. Considering the definition of the effective price elasticity in equation (5) of the text, this assures that this elasticity cannot be negative, for any exit rate. Another implication of the border conditions is that (A.3.a) and (A.3.b) together require that $A^*(\infty)$ be bounded. This may be interpreted as a requirement that marketing messages and price reductions should not be too effective in attracting new clients.

Second order conditions

The second order condition for program (2) is that second variations of the controls p, M and m over H be a negative semidefinite matrix. The diagonal elements must be negative:

(A.4.a) $H_{pp} < 0$ => $\lambda \cdot [\partial^2 E/\partial p^2 - A \cdot \partial^2\phi/\partial p^2] < 0$

(A.4.b) $H_{MM} < 0$ => $\lambda \cdot \partial^2 E/\partial M^2 < 0$

(A.4.c) $H_{mm} < 0$ => $\lambda \cdot [(\partial^2 E/\partial^2 m) + (^{-}\partial^2\phi/\partial^2 m)\cdot A] < 0$

to which must be added the other conditions (not shown).

Optimal policies in the transition path

Replacing (A.2a') in (A.2b) we find that:

$$\text{(A.5)} \qquad \frac{p^* - c_a - k_m m + d l/dt}{r + f^*} \cdot \frac{\partial E}{\partial M}(M^*(t)) = c_m$$

Equation (A.5) shows how optimal marketing policy may change from the short to the long run. In the short run, the marginal benefit of a message includes the capital gain dλ/dt. For example, if the firm believes that marginal production costs mill fall over time, equation (A.2a') shows that the value of a client wil increase steadily over time. This generates an expected capital gain, and the optimal path takes it into account by recommending faster investment in messages M. A similar argument holds for m*(t).

Regarding pricing policy, insertion of (A.2b) in (A.2a) yields:

$$\text{(A.6)} \qquad \frac{p^* - c_a - k_m m + d l / dt}{p^*} = \frac{1}{\left[h^E \cdot \left(\frac{E}{Af}\right) + h^f\right] \cdot \left(\frac{f}{r+f}\right)}$$

Equation (A.6) shows that in the short run the optimal pricing policy takes into account the capital gain already discussed and the ratio between the "entry rate" E/A and the exit rate ϕ. A positive capital gain induces the firm to reduce prices below long run levels to speed up entry and reduce exit and appropriate the coming capital gain. In addition, when the entry rate exceeds the exit rate, the elasticity of entries to price receives a higher weighting than the elasticity of exits to price.

In the steady state -long run - the conditions dλ/dt =0 and ϕ = E/A hold, so (A.6) settles into equation (5) of the text. Equation (A.5) settles into equation (3) of the text. Therefore, (3), (4)and (5) are valid in the steady state only.

Uniqueness of the optimal path

If program (2) is consistent with economic experience, the optimal path should be unique. To check uniqueness we use a phase diagram in some control variable - usually price - and the state variable, the number of clients A. Depending on the signs of the slopes of curves dA/dt = 0 and dp/dt =0 in this diagram, the number of paths that meet the first order conditions can be zero, one or many. In particular, if curve dA/dt = 0 has a negative slope and curve dp/dt =0 has a positive slope, then the number of paths that meet the first order conditions is one, and uniqueness is proved.

Program (2) includes three control variables: p, M y m. To simplify the proof of uniqueness, we suppress controls M and m, assuming they take fixed values. This includes the case where c_m and k_m are high enough to make M* = 0 and m* = 0.

Equation (A.2e) indicates that: dA/dt = E(p) - ϕ(p)·A Thus, there exists a curve where dA/dt = 0 in plane (p,A). The slope of this curve is obtained by differentiating (A.2e) totally. The result is:

$$\text{(A.7)} \qquad \left.\frac{dp}{dA}\right|_{\frac{dA}{dt}=0} = \frac{-f}{-E_p + f_p A}$$

Recall that equation (A.2.b) implies that the denominator is positive because A/λ > 0. As ϕ > 0, equation (A.7) proves that the slope of curve dA/dt = 0 is negative throughout.

To find the curve dp/dt =0 we use the other first order conditions. First, we differentiate (A.2b) with respect to time. Then we replace the term in d λ/dt using equation (A.2.a). We also replace the terms in dA/dt using equation (A.2.e). Finally the terms in λ are replaced using equation (A.2.b). The result is:

(A.8) $$\frac{dp}{dt} = \left(\frac{E}{A} - f\right)\frac{(-E_p)}{\left(-E_{pp} + Af_{pp}\right)} + \left(\frac{(p - c_a - k_m m)}{A/(-E_p + Af_p)} - (r + f)\right)\left[\frac{-E_p + Af_p}{E_{pp} + A\cdot(-f_{pp})}\right]$$

This shows the evolution of price over time along optimal paths. The curve dp/dt =0 is obtained setting the left-hand side of (A.8) to zero. The resulting equation is a relationship between p and A:

(A.9) $$p = (c_a + k_m m) + \frac{(r + f)\cdot A}{-E_p + Af_p} + \frac{(f\cdot A - E)(-E_p)}{(-E_p + Af_p)^2}$$ (curve dp/dt=0)

Recall that the expression in both denominators is positive by (A.2b). To find the slope of this relationship in plane (p,A), we differentiate (A.9) totally with respect to p and A. The result is a complex expression, but it simplifies radically at the point(s) where this relationship crosses the curve dA/dt = 0, because in that crossing dA/dt = E - ϕA = 0. We find that in every crossing point between the curves dp/dt =0 and dA/dt =0, the former curve has the slope:

(A.9) $$\left.\frac{dp}{dA}\right|_{\frac{dp}{dt}=0} = \frac{-E_p}{A}\cdot\frac{(r + 2f / r + f)}{(-E_{pp} + f_{pp}A)} > 0 \qquad where \ \ \frac{dA}{dt} = 0 \ \ is \ \ crossed\,.$$

We know that (A.4a) requires the numerator to be positive. As $E_p < 0$ by assumption, this expression is positive. It is possible that the curve dp/dt =0 may have a negative slope in other regions of plane (p,A), but locally, where the two curves cross, its slope is always positive. We have no ruled out the possibility of multiple crossings.

Summing up, the curves dA/dt =0 and dp/dt =0 have slope with the signs shown in Figure 1, at least locally near the steady state. The laws of movement in this figure imply that the optimal path is unique. The figure shows two optimal paths because we have not determined whether the optimal path is such that prices increase or fall when approaching the steady state from a number of clients below the one of the steady state.

In the more general case where the firm controls M and m in addition to p, much more elaborate calculations are needed to prove uniqueness. Generally speaking, uniqueness requires that the cross derivative E_{pM} not to be too large and positive. This is a natural assumption, because the second order conditions limit the size of these cross derivatives.