CCHE CCES
Canadian Centre for Health Economics
Centre canadien en économie de la santé

Working Paper Series
Document de travail de la série

PATENT EXPIRATION AND COMPETITION: A DYNAMIC LIMIT PRICE MODEL

Anastasios I. Papanastasiou

Working Paper No: 2014-10

www.canadiancentreforhealtheconomics.ca

May 1, 2014

Canadian Centre for Health Economics
Centre canadien en économie de la santé
155 College Street
Toronto, Ontario

CCHE/CCES Working Paper No. 2014-10
May 1, 2014

Patent Expiration and Competition: A dynamic limit price model

Abstract

We develop a dynamic model to explore the optimal pricing strategy of a monopolist that faces potential market entry at a given point in time. By engaging in promotional activities, the dominant firm may increase future demand for the product, while by charging below a limit price it can prevent competition from entering the market. Our analysis suggests that the optimal path for price and advertisement depends on the price elasticity of demand and the duration of monopoly life. Relating our model to the market for pharmaceuticals, we establish conditions that would give rise to a Generics Competition Paradox (GCP) and discuss how these conditions are linked to the existing theories that attempt to explain the GCP.

JEL Classification: D21; D42; I11; L12; C61

Key words: monopoly, generic competition, brand-name drugs, limit price, price elasticity of demand

Corresponding Author:

Anastasios I. Papanastasiou*
Department of Economics & Finance
University of Guelph
50 Stone Road East
Guelph, N1G 2W1
Ontario, Canada
Email: apapanas@uoguelph.ca

I. Introduction

Manufacturers of prescription drugs are granted a period of patent protection upon the introduction of a new product. For a period of time, patent protection implies immunity from market competition thus allowing the manufacturer to act as a monopolist and establish their brand-name in association to the new chemical substance. The expected monopoly profits encourage firms to engage in risky investments (e.g. research, testing etc.) that would possibly lead to the development of new beneficial drugs. After patent expiration and in order to minimize the monopoly costs to society, manufacturers of generic drugs may enter the market and brand-name firm eventually faces rival competition from firms that market virtually the same product. Typical intuition may suggest that prices tend to fall when markets become more competitive - for example in some limit price models it is optimal for the firm to charge below the monopoly price when the market opens to competition in order to deter rival entry. If producers' price is the only cost to consumers and the products in the market are close substitutes then, one may expect prices to converge and firms with high prices to receive a market share of zero. However there is a substantial amount of evidence suggesting that this is not the case in the prescription drug market.

Grabowski and Vernon (1992) observe that in most cases out of 18 drug products that had received patent protection in the U.S. market, brand-name firms would continue to increase their price after patent expiration despite the fact that their generic competitors entered the market quoting a much lower price for their product. They find that it takes about two years after patent expiration for the typical product in the sample to lose half of the market to competition and that the number of competitors in each market is statistically related with market profitability. Caves *et al.* find no evidence of limit pricing and estimate that generic drugs capture roughly 25% of the market when the ratio of generic to branded price is 0.45. The contradiction between the empirical evidence and economists' typical intuition became known as the 'Generics Competition Paradox' (GCP) (Scherer, 1993).

More recent evidence documenting this paradox in the U.S. markets include, e.g., Saha *et al.* (2006) and Regan (2007). A similar phenomenon has been observed in the Canadian and E.U. drug prescription market (e.g. Vandoros and Kanavos (2012), Kanavos *et al.* (2008)).

Using a two period stackelberg model with several entrants, Frank and Salkever (1992) show that it may be optimal for the brand-name firm to increase its price as generic drugs enter the market if rival entry causes demand for the brand-name product to become less elastic. This result assumes that the demand consists of some consumers who are loyal to the brand-name firm so that the latter is able to maintain its monopoly power over the former after patent expiration: only the segment of the demand that consists of consumers who are cross price sensitive may switch to purchasing the generic drugs after patent expires.

Bhattacharya and Vogt (2002) propose a dynamic theory where they assume a one-time effect of patent expiration on brand-name demand. In their framework, the profit maximizing firm charges a higher price for its product after patent expiration under the assumption that rival entry makes demand more elastic which is in contradiction to the result and the assumption of Frank and Salkever.

Other explanations for the GCP focus on the role of third party institutes such as insurance coverage. The GCP was documented only in the prescription drug market where out of pocket expenditure may significantly vary among consumers depending on their type of insurance. Ferrara and Kong, 2008 show that it may be optimal for the firm to increase its price after rival entry under some conditions when the demand consists of groups with different insurance coverage.

In our approach, we assume that rival entry causes brand-name demand to become less elastic. We expand upon the aforementioned literature by assuming that the effect of rival entry on brand-name demand is continuous in time. During the monopoly period we assume that consumer loyalty can be represented by a stock variable, its growth depending on the amount of current advertisment exposure. As in Bhattacharya and Vogt, time is modelled continuously and the problem of the firm during both the pre-entry and post-entry phase is to choose the profit maximizing path for advertisement and price subject to the constraint that consumer loyalty obeys a given law of motion.

II. Model

A. Assumptions

The firm faces a demand of the form $Q(p(t), F(t))$ where t, p and F are non-negative real variables denoting time and product price respectively, F is a measure of the consumers' loyalty with the product.

We assume that Q is twice continuously differentiable in (p, F), with finite partial derivatives and satisfies $(i) - (vi)$:

$(i).\ \frac{\partial Q(p,F)}{\partial p} < 0,\quad (ii).\ \frac{\partial^2 Q(p,F)}{\partial p^2} = 0,\quad (iii).\ \frac{\partial Q(p,F)}{\partial F} > 0,\quad (iv).\ \frac{\partial^2 Q(p,F)}{\partial F^2} < 0,$

$(v).\ \frac{\partial^2 Q(p,F)}{\partial p \partial F} < 0$ and $(vi).\ Q(p, 0) > 0$ for some $p > 0$.

Assumptions (i) and (ii) imply that quantity demanded decreases at a constant rate as price increases , while assumptions (iii) and (iv) imply that quantity demanded increases at a decreasing rate as consumer loyalty increases.

Through assumption (v) an increase in F causes quantity demanded to fall more, for a given increase in price. That is, an increase in F rotates the linear demand curve in (Q, p) making it more price elastic.

For $t \in [0, T]$, which we shall call the first phase, the firm acts as a monopolist. During this phase familiarity builds up according to the equation of motion:

$$\dot{F}(t) = g(A(t)) - \delta F(t)$$

where the variable A represents the real level of advertising exposure, the parameter $\delta \in (0,1)$ is the rate of loyalty depreciation which is assumed to be constant over time, and g has the standard properties of a production function: $g_A > 0$, $g_{AA} < 0$ and $g(0) = 0$. Thus, the firm can not directly control the level of $F(t)$ but by choosing the level of $A(t)$ it affects the future values of the former.

We assume that the firm can obtain units of A at a constant price per unit v. Since g is an increasing function an additional unit of $A(t)$ causes $\dot{F}(t)$ to increase, as a result future quantity demanded rises (assumption (iii)) and demand becomes more price elastic (assumption (v)).

Intuitively, we are looking at a population of potential consumers of a fixed size - the population with the condition that the drug is aimed at. With no advertisement, $F = 0$ and for some $p > 0$ where $Q > 0$ there exist some consumers that choose not to buy the drug, presumably having done the calculation and decided that the price was not matched by the marginal health benefit. An increase in advertisement exposure will attract some of the consumers who are more price sensitive, or have a lower willingness to pay, than the ones that already consume the product. This increases market size by bringing in a more price elastic segment of the market and as a result demand overally becomes more elastic.

After time T, which we will refer to as the second phase or the competition phase, there are no barriers to entry and the law of motion for F becomes:

$$\dot{F}(t) = g(A(t)) - \delta F(t) - \gamma(p(t) - \overline{p})$$

The term $\gamma(p(t) - \overline{p})$ represents the rate of entry of rival producers. If the number of competitors in each market is related with market profitability and if rival competitors view current product price as a proxy for market profitability then the rate of entry of rival producers must be a function of current product price (Gaskins, 1970). We assume that rival entry is a linear function of current product price and define $\overline{p}$ as the price at which net entry is equal to zero with γ being a positive response coefficient. Pricing above the limit price $\overline{p}$ causes positive rival entry at time t: the increase in the term $\gamma(p(t)-\overline{p})$ has no effect on $Q(p(t), F(t))$ at time t, but by decreasing $\dot{F}(t)$ it reduces the dominant firms' demand in the future through (iii), ceteris paribus. Moreover, the reduction in F makes demand more price inelastic through assumption (v). As rival competitors enter the market and F is reduced it is the more price sensitive individuals who leave, making the demand less elastic and shifting it towards the origin.

In the second phase F can no longer be interpreted as consumer loyalty - it is the channel through which advertisement *and* rival sales affect future demand.

The problem of the firm is to choose the path for price and advertisement exposure that will maximize the present value of its lifetime profits. We further assume that future income is discounted at a constant rate r and that c, the average total cost of production, is also constant and that $c < \overline{p}$.

B. Necessary Conditions

Formally, the problem can be written as

$$\max_{(p(t),A(t))} \int_0^{\infty} [(p(t)-c)Q(p(t),F(t)) - vA(t)]e^{-rt}dt$$

subject to

$$\dot{F}(t) = g(A(t)) - \delta F(t) - \gamma(p(t)-\overline{p}) \qquad (1)$$

$$\gamma = 0 \; if \; t \in [0,T], \; otherwise \; \gamma > 0$$

$$F(0) = 0$$

The current value Hamiltonian associated with the above maximization problem can be expressed as

$$H(p, F, A, \psi) = (p - c)Q(p, F) - vA + \psi[g(A) - \gamma(p - \overline{p}) - \delta F]$$

where the time argument is omitted for notational simplicity. In the language of optimal control, p and A are control variables, F is the state variable and ψ is the co-state variable. In economic terms, ψ is the value associated with a marginal increase in F or the current value shadow price of F.

The necessary conditions are:

$$0 = \frac{\partial H}{\partial p^*} \Rightarrow$$

$$0 = Q(p^*, F) + p^* Q_p(p^*, F) - cQ_p(p^*, F) - \gamma\psi \tag{2}$$

$$0 = \frac{\partial H}{\partial A^*} \Rightarrow$$

$$0 = -v + \psi g_A(A^*) \tag{3}$$

$$\dot{\psi} = r\psi - H_F \Rightarrow$$

$$\dot{\psi} = (r+\delta)\psi - (p-c)\,Q_F \tag{4}$$

$$\dot{F} = g\,(A) - \gamma(p-\overline{p}) - \delta F \tag{5}$$

and must hold for all t, in addition to the transversality condition

$$e^{-rt}\psi^* = \frac{\partial\varphi}{\partial F^*} \tag{6}$$

which must be satisfied at $t = T$,where $\varphi = \max\limits_{(p(t),A(t))} \int_T^\infty [(p(t)-c)Q(p(t),F(t)) - vA(t)]e^{-rt}dt$ subject to (1) with $\gamma > 0$ and $F(T)$ given. Condition (6) can be thought as a terminal condition for phase one and prevents the optimal value of the co-state from changing discontinuously at time T.

Condition (2) can be rewritten as

$$cQ_p(p^*, F) + \gamma\psi = Q(p^*, F) + p^*Q_p(p^*, F)$$

and during the first phase simplifies to

$$cQ_p(p^*, F) = Q(p^*, F) + p^*Q_p(p^*, F)$$

The above conditions implicitly define the optimal price $p^*(F,\psi)$ as a function of the state and co-state variables. In the first period $\gamma = 0$ and the optimal price is a function of the state variable only since the term $\gamma\psi$ vanishes. The economic intuition behind these conditions is that the optimal pricing policy $p^*(t)$ must equate marginal cost to marginal revenue at every instant. During the second period, the term $\gamma\psi$ reflects the additional costs that stem from increased competition: pricing above the limit price will result in positive rival entry which will decrease F in the future. This additional cost to the firm is equal to the change in F[1] times the shadow price of F.

Similarly condition (3) implicitly defines the optimal amount of advertisement exposure $A^*(\psi)$ as a function of the co- state variable and has the same economic interpretation: the marginal cost of advertisement must be equal to the marginal benefit. In addition, (3)implies that the current value[2] shadow price of F is always positive.

Substituting conditions (2) and (3) into the law of motion for F and ψ yields a pair of autonomous differential equations:

(7) $$\dot{\psi} = (r+\delta)\psi - [p^*(F,\psi) - c]Q_F(p^*, F)$$

(8) $$\dot{F} = g(A^*(\psi)) - \gamma[p^*(F,\psi) - \overline{p}] - \delta F$$

From (2), we can calculate

$$p_F^* = \frac{Q_F + (p-c)\,Q_{pF}}{-2Q_p}$$

and

$$p_\psi^* = \frac{\gamma}{2Q_p} < 0$$

We will assume that $Q_F > -Q_{pF}$ through most of F so that additional units of loyalty will increase the equilibrium price. Intuitively, a decrease in consumer loyalty during the first phase causes demand to fall and also become more price inelastic. If the former effect dominates the latter then the firm will decrease its price in order to re capture the demand that was lost due to a decrease in F.

From (3) we obtain

$$A^*_\psi = -\frac{g_A(A)}{\psi g_{AA}(A)} > 0$$

As expected the optimal amount of real advertisement exposure increases as F becomes more valuable.

The stationary locus for $\psi(p^*, F(A^*))$ is defined as

$$0 = (r+\delta)\psi - [p^*(F,\psi) - c]Q_F(p^*, F) \tag{9}$$

and it follows by the implicit function theorem that

$$\frac{d\psi}{dF}\,|_{\dot{\psi}=0} = \frac{p^*_F Q_F + (p^* - c)(Q_{FF} + Q_{Fp^*}p^*_F)}{(r+\delta) - [(p^* - c)Q_{Fp^*} + Q_F]\,p^*_\psi} \tag{10}$$

In phase 1,the right hand side of the above expression is positive for small values of F and becomes negative as F increases. Thus, $\dot{\psi} = 0$ is increasing in F for small values of F and decreasing for large values and the locus must be concave in (ψ, F) (see appendix).

In phase 2, the additional term in the denominator $[(p^* - c)Q_{Fp^*} - Q_F]\,p^*_\psi$ is positive and causes ψ to change less rapidly along $\dot{\psi} = 0$, as F increases. Assuming that $(r+\delta) - p^*_\psi Q_F > (p^* - c)Q_{Fp^*}$ the locus is first increasing and then decreasing in F, as in phase 1, otherwise the first order properties in F are reversed.

In addition, during phase one, equation (9) requires the value of ψ to be positive at $F = 0$, say $\widetilde{\psi}$. In phase 2, the value of ψ is positive and less than $\widetilde{\psi}$ if some loyalty has accumulated up to time T, since Q_F will be smaller in magnitude through assumption (iv).

We now turn our attention to the slope of the stationary locus for the state variable:

$$0 = g(A^*(\psi)) - \gamma[p^*(F, \psi) - \overline{p}] - \delta F \tag{11}$$

For this locus we have

$$\frac{d\psi}{dF} \mid_{\dot{F}=0} = \frac{[\gamma p_F^* + \delta]}{[g_A A_\psi^* - \gamma p_\psi^*]} \tag{12}$$

Therefore this locus is increasing in F during the first period and also during the second, assuming $g_A A_\psi^* > \gamma p_\psi^*$.

If the value of ψ at $F = 0$ is less than $\widetilde{\psi}$ then, by continuity, the two loci intersect at (F^*, ψ^*) where $\dot{\psi}(F^*, \psi^*)$ and $\dot{F}(F^*, \psi^*)$ are both equal to zero.

In addition

$$\begin{aligned} \frac{d^2\psi}{dF^2} \mid_{\dot{F}=0} &= \left[\gamma p_{FF}^*(g_A A_\psi^* - \gamma p_\psi^*) - (-\gamma p_{\psi F}^*)(\gamma p_F^* + \delta)\right] \\ &\quad * (g_A A_\psi^* - \gamma p_\psi^*)^{-2} \end{aligned}$$

The above expression is negative which means that the stationary locus for F is concave in (F, ψ).

Since $p_\psi = 0$ in phase 1 this expression becomes

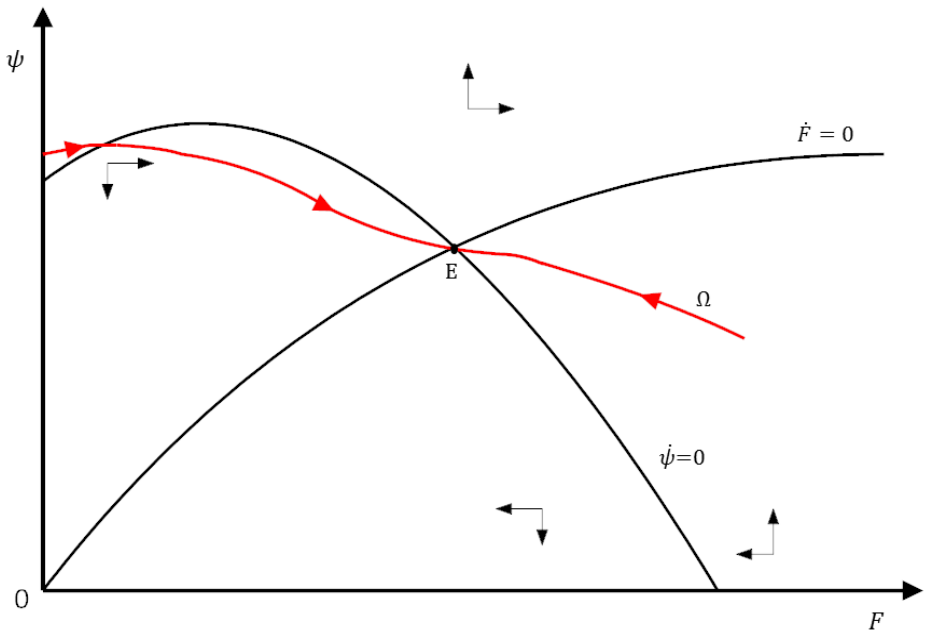

Figure 1: Stationary loci for the state and co − state variables

$$\frac{d^2\psi}{dF^2} = \frac{\gamma p^*_{FF} g_A A^*_\psi}{(g_A A^*_\psi)^2} = 0$$

However, with F on the horizontal axis the formula for the curvature of this locus during phase one is given by

$$\frac{d^2F}{d\psi^2} = \frac{g_{AA} A^{*2}_\psi + A^*_{\psi\psi} g_A}{\delta}$$

which is negative since $A^*_{\psi\psi} < 0$.

Therefore the stationary locus for F is concave in (F, ψ) during both phases.

Figure 1 summarizes our findings on the stationary loci. The two loci divide the (ψ, F) space into four regions with different dynamics, represented by the phase arrows, that determine the time path of the system in each region. There is a unique trajectory that reaches the steady state E where both F and ψ are not growing over time, represented by the stable path Ω. That is for all initial pairs $(\psi, F) \in \Omega$ the system converges to E. Any other trajectory going through $(\psi, F) \notin \Omega$ eventually diverges from E. After time T the loci shift but have the same first and second order properties under our assumptions. The dynamics of the space change accordingly so that the phase two dynamics are the same as in Fig.1 with respect to the position of the phase two stationary loci. We will now examine the direction of the shift of the loci after patent expiration.

C. The Effect of Rival Entry

Totally differentiating (9) yields

$$\frac{d\psi}{d\gamma}\Big|_{\dot{\psi}=0} = \frac{p^*_\gamma[Q_{Fp}(p^*-c)+Q_F]}{(r+\delta)-[(p^*-c)\,Q_{Fp}+Q_F]\,p^*_\psi} \tag{13}$$

The above expression is positive if

$$\frac{(r+\delta)}{p^*_\psi} - Q_F < (p^*-c)Q_{Fp} < -Q_F$$

and negative if

$$(p^*-c)Q_{Fp} > -Q_F$$

or

$$(p^*-c)Q_{Fp} > \frac{(r+\delta)}{p^*_\psi} - Q_F$$

From the first order condition for p we can calculate

$$\frac{dp^*}{d\gamma} = \frac{\psi}{2Q_p} < 0$$

Recall that we have assumed that $Q_F > -(p-c)Q_{pF}$ through most of F in order for $\frac{dp^*}{dF} > 0$.

Thus, $(p^* - c)Q_{Fp} > -Q_F$ which implies that the post entry stationary locus for ψ will be *below* the respective pre entry stationary locus.

The stationary locus for F is given by $0 = g(A^*(\psi)) - \gamma[p^*(F,\psi) - \overline{p}] - \delta F$.

Therefore

$$\frac{d\psi}{d\gamma}\Big|_{\dot{F}=0} = \frac{(p^* - \overline{p}) + \gamma p^*_\gamma}{g_A A^*_\psi - \gamma p^*_\psi}$$

The denominator is positive so the sign of the above expression will depend on the sign of the numerator, specifically if $(p^* - \overline{p}) > -\gamma p^*_\gamma$ (less than), then the stationary locus for F post entry is above (below) the respective locus in the pre entry phase. So there are two cases to consider:

Case 1: $(p^* - \overline{p}) > -\gamma p^*_\gamma$

This case is depicted in Figure 2 with a phase diagram where we have drawn the stationary loci for both phases. During phase one, the loci intersect at E_1 while during phase two at E_2. Working backwards we are able to sketch the optimal trajectory of the system. Since the second phase has an infinite horizon, when the system is controlled optimally it approaches the steady state E_2 asymptotically. There is a unique locus of points in (ψ, F) along which the system can reach E_2, denoted by ω_2. Thus $(\psi^*, F^*) \in \omega_2$ for all $t > T$. The transversality condition for our two phase problem (6) requires ψ^* to be continuous at time T, that is, the value of ψ^* at the beginning of the second phase is $\psi^*(T)$. For this to be satisfied, the optimal trajectory must approach a point in ω_2 for some time up to time T . Now consider the first phase of the problem. At $t = 0$ the value of familiarity is zero so the planner picks a point on the vertical axis of Figure 2. The optimal trajectory must reach reach a point in ω_2 at exactly time T without violating the direction of the phase

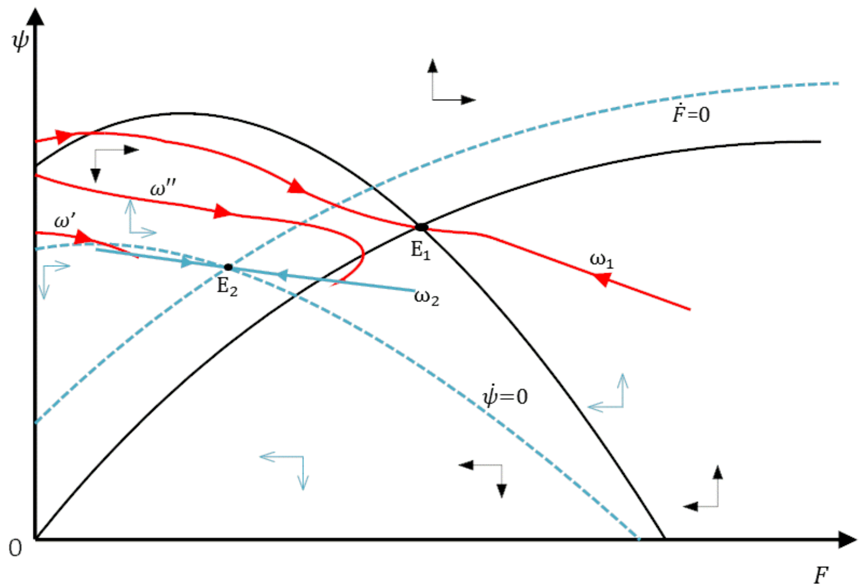

Figure 2: Optimal Trajectories in Case 1

arrows. Thus, the set of optimal trajectory candidates are all the trajectories that satisfy $F(0) = 0$ and reach a point in ω_2 at time T. The length of the first phase (i.e. the value of T) determines the optimal trajectory uniquely. In fig.2, ω' and ω'' are both optimal trajectories in the sense that they satisfy the optimality conditions $(1) - (4)$ given $F(0) = 0$. If T is small then the unique optimal trajectory is ω' and if T is large then the unique optimal trajectory is ω''. Any other trajectory fails to reach a point in ω_2 at time T and satisfy the transversality condition. As $T \to \infty$ the optimal trajectory reaches a neighborhood close to E_1 where the growth rate of the state and co-state variables are infinitely small.

Assume that T is such that ω'' is the optimal trajectory. Then time can be decomposed into three different time periods in which the control variables change monotonically:

Time period I : From $T = 0$ and until the system crosses the phase one stationary locus for F , ψ decreases while F increases over time. During this time period A^* is falling while p^* is rising and $\dot{F} \geq 0$, the equality holds for the instant that the trajectory crosses the $\dot{F} = 0$

locus, say t'.

Time period II : After the system crosses the stationary locus for F in phase 1, ψ decreases over time but now F is also falling over time. Throughout the duration of this period, A^* is falling and p^* may either fall or rise over time depending on the magnitude of p^*_F relatively to that of p^*_ψ.

Time Period III : After the end of Phase 1 the system asymptotically approaches E_2 and $(\psi^*, F^*) \in \omega_2$ for the whole duration of phase 2. During this period, ψ is rising and F is falling. This causes p^* to fall and A^* to rise over time.

Therefore in Case one the optimal solution directs price to fall after rival competitors enter the market. This result is not a new one: it implies that prices tend to fall when markets become more competitive.

However, when T is small then the optimal trajectory is ω' since it is the unique trajectory that reaches ω_2 at exactly time T. If this is the case then the system approaches E_2 from the left during phase two and the optimal values for price is increasing in this case and advertisement is falling during the whole duration of the problem.

Case 2: $(p^* - \overline{p}) < -\gamma p^*_\gamma$

This case is depicted in Figure 3. In this case the optimal path of price is rising post-patent expiration *for all* optimal trajectories with finite T. This is a result in contradiction with many models of competition. If the duration of the first phase is long enough then a trajectory such as ω'' or ω''' is optimal. For small values of T trajectories such as ω' are optimal. Any trajectory that begins above the phase one stable arm, such as ω_ε in fig.3, can not be optimal since it diverges from E_2 as $T \to \infty$.

In both cases 1 and 2, the optimal price is increasing and F is growing for $t \in [0, t')$. This is also the case for optimal trajectories that begin above the stationary locus for ψ such as ω''' in figure 3. After time t' and up to T the system is approaching ω_2 and A^* is

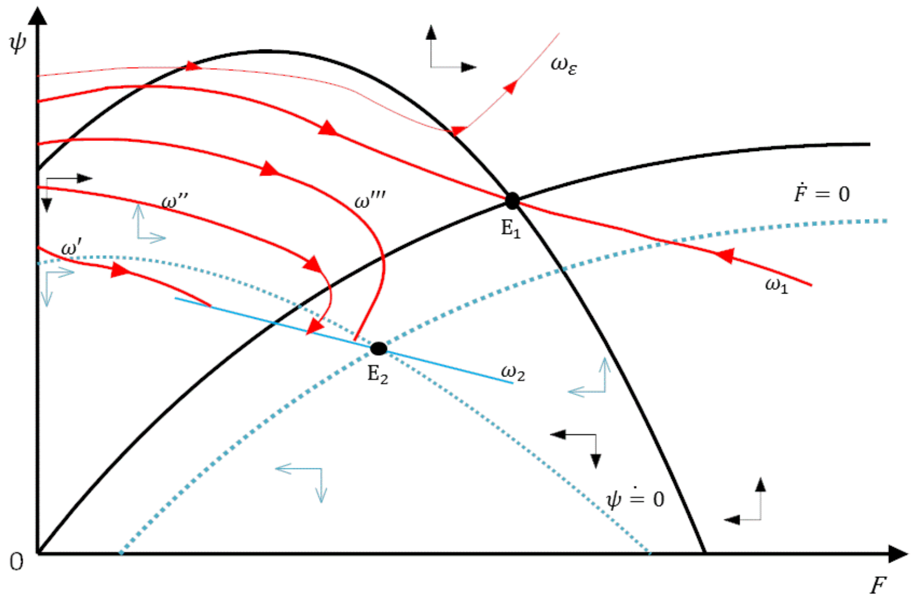

Figure 3: Optimal Trajectories in Case 2

still decreasing while the optimal price is increasing if $p_F^* > p_\psi^*$ and decreasing otherwise. However if the duration of the first phase is short then time period II does not occur because there is not enough time for the optimal trajectory to cross the stationary locus for F during the monopoly phase. This implies that at the beginning of the problem, when $F = 0$ it is optimal for the firm to start with a high value of real promotional exposure[3] and then gradually reduce it up to time t'. As consumer loyalty accumulates, the firm finds it profitable to increase its price. The continuous decrease in A^* causes loyalty to start growing negatively after t' and since $\frac{dA^*}{dt}$ remains negative throughout $t \in (t', T]$, loyalty will decrease at a faster rate as time approaches patent expiration. After the time that the growth rate of loyalty becomes negative[4], it is optimal for the firm to charge a higher price for its product if the total effect of ψ is greater than the total effect of F,on p^*. Intuitively, the firm would like to settle at E_1 if the monopoly phase would last forever. As patent expiration approaches and advertisement is falling, consumer loyalty starts to grow negatively. This has two results, it reduces quantity demanded and it makes demand more price inelastic. The firm is still a monopolist in the market therefore the decrease in the price elasticity of demand would have a positive effect on the optimal price but since the firm now faces imminent rival entry, any increases in price would result in a higher rate of rival entry after patent expiration and therefore greater reductions in quantity demanded and own-price elasticity. Depending on the relative magnitude of our exogenous parameters, the firm may choose to lose some of the more price sensitive consumers by increasing its price as patent expiration approaches.

When the horizon of post-patent competition is infinite our model predicts that the optimal price may either rise or fall depending on the magnitude of $(p^* - \overline{p})$ relatively to $-\gamma p^*_\gamma$. If the optimal price is close to the limit price and Q_p is not too large, then $(p^* - \overline{p}) < -\gamma \frac{\psi}{2Q_p}$ and the optimal price is increasing post patent expiration for all finite values of T. On the other hand, if the limit price is too low and Q_p is large, then $(p^* - \overline{p}) > -\gamma \frac{\psi}{2Q_p}$ and the optimal price is decreasing post patent expiration.

If the first period is short then the optimal trajectory may approach ω_2 without crossing the stationary locus for F, so time period II may not occur; in this case if $(p^* - \overline{p}) > -\gamma \frac{\psi}{2Q_p}$ then p^* may be decreasing *or increasing* after T , whereas if $(p^* - \overline{p}) < -\gamma \frac{\psi}{2Q_p}$ then $\frac{dp^*}{dt} < 0$ for all $t > T$. Finally, as $T \to 0$, $(\psi^*, F^*) \in \omega_2$ at $t = 0$ and $\frac{dp^*}{dt} > 0$ for all $t > T$.

III. Conclusion

Our model predicts that the optimal price tends to increase after patent expiration in a number of circumstances. We have shown that the duration of the first phase is crucial in determining the optimal behavior of the firm. If the duration of the patent is very short it is always optimal for the firm to increase its price after patent expiration. The pharmaceutics R&D industry is highly competitive and as a result, manufacturers patent new chemical substances early in the trial stages. While patent protection is twenty years, many novel drugs end up with a monopoly life of eight years as a result of early patenting. Our model suggests that short monopoly life is a possible explanation for the GCP.

On the other hand, if the first period is long then the it may be optimal for the firm to increase its price as patent expiration approaches or after patent expiration. In the pre-entry phase, the firm may increase its price in adjustment towards the new equilibrium where it sells to the more price inelastic part of the demand (i.e. the consumers with higher willingness to pay) exclusively. In the market for prescription pharmaceuticals such a pricing policy would imply that some of the patients that were receiving medication would have to discontinue their treatment until generic drugs entered the market.

A Appendix

The curvature of $\dot{\psi} = 0$ can be determined from

$$
\begin{aligned}
\frac{d^2\psi}{dF^2} \mid_{\dot{\psi}=0} = & \{[p^*_{FF}Q_F + 2(Q_{FF} + Q_{Fp})p^*_F + (p^* - c)(Q_{FFF} + Q_{FFp}p^*_F + Q_{FpF}p^*_F + Q_{Fpp}p^{*^2}_F + p^*_{FF}Q_{Fp}) \\
& + p^*_F(Q_{FF} + Q_{Fp}p^*_F)] \\
& * \left[(r+\delta) - [(p^* - c)Q_{Fp} + Q_F]p^*_\psi\right] \\
& + \left[p^*_{\psi F}[(p^* - c)Q_{Fp} + Q_F] + [(p^* - c)(Q_{Fpp}p^*_F + Q_{FpF}) + p^*_F Q_{Fp} + (Q_{FF} + Q_{Fp}p^*_F)]p^*_\psi\right] \\
& * [p^*_F Q_F + (p^* - c)(Q_{FF} + Q_{Fp}p^*_F)]\} \\
& * \left((r+\delta) - [(p^* - c)Q_{Fp^*} - Q_F]\, p^*_\psi\right)^{-2}
\end{aligned}
$$

During phase one, $\gamma = 0$ and assuming that the third derivatives of Q are zero the above expression becomes

$$\frac{d^2\psi}{dF^2} \quad | \quad _{\dot{\psi}=0} = [p^*_{FF} Q_F + 2(Q_{FF} + Q_{Fp})p^*_F + (p^* - c)(p^*_{FF} Q_{Fp})$$
$$+ p^*_F (Q_{FF} + Q_{Fp} p^*_F)] * (r + \delta)^{-1}$$

Differentiating the first order condition for price:

$$\begin{aligned} p^*_{FF} &= \{[Q_{FF} + (p-c)Q_{pFF}](-2Q_p) + 2Q_{pF}[Q_F + (p-c)Q_{pF}]\} * (-2Q_P)^{-2} \\ &= [Q_{FF} + (p-c)Q_{pFF}](-2Q_p) + 2Q_{pF}Q_F + 2Q^2_{pF}(p-c)] * (-2Q_P)^{-2} \end{aligned}$$

thus $p^*_{FF} < 0$ if $[Q_{FF} + (p-c)Q_{pFF}](-2Q_p) + 2Q_{pF}Q_F > 2Q^2_{pF}(p-c)$,

and

$$p^*_{\psi F} = -\frac{\gamma}{2(Q_{pF})^2} < 0$$

Therefore, $\frac{d^2\psi}{dF^2} < 0$ during the first phase.

REFERENCES

Jayanta Bhattacharya and William B. Vogt (2002), "A Simple Model of Pharmaceutical Price Dynamics", *Journal of Law and Economics* 46 (2)*:* 599-626

Richard E. Caves, Michael D. Whinston, Mark A. Hurwitz, Ariel Pakes, Peter Temin (1991), "Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry" *Brookings Papers on Economic Activity: Microeconomics,* 1991, 1-66

Ida Ferrara and Ying Kong (2008), "Can health insurance explain the generic competition paradox?", *Economics Letters* 101*:* 48-52

Richard G. Frank and David S. Salkever (1992), "Pricing, Patent Loss and the Market for Pharmaceuticals " *Southern Economic Journal* 59 (2): 165-79

Gaskins, Darius W. (1971), "Dynamic Limit Pricing: Optimal Pricing Under Threat of Entry", *Journal of Economic Theory* 3 : 306–322.

Henry G. Grabowski and John M. Vernon (1992), "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act" *Journal of Law and Economics* 35: 331-350

Panos Kanavos, Joan Costa-Font and Elizabeth Seeley (2008), "Competition in off-patent drug markets: Issues, regulation and evidence", *Economic Policy* 499–544

Ying Kong (2009), "Competition between brand-name and generics - Analysis on pricing of brand-name pharmaceutical", *Health Economics* 18: 591-606

Atanu Saha, Henry Grabowski, Howard Birnbaum, Paul Greenberg, Oded Bizan (2006), "Generic Competition in the U.S. Pharmaceutical Industry", *International Journal of the Economics of Business 1*3(1): 15-38

Scherer F.M. (1993) "Pricing, profits, and technological progress in the pharmaceutical industry" *Journal of Economic Perspectives* 7 (3), 95–115

Tracy L. Regan (2008), "Generic entry, price competition and market segmentation in the prescription drug market" *International Journal of Industrial Organization* 26: 930-948

Sotiris Vandoros and Panos Kanavos (2012), "The generics paradox revisited: empirical evidence from regulated markets", *Applied Economics* 45:22, 3230-3239

Notes

*Address for correspondence: Department of Economics, 50 Stone Road E, Guelph, N1G 2W1, Ontario, Canada. E-mail: apapanas@uoguelph.ca

[1]The price induced change in the growth of F is equal to $-\gamma$ when price is increased marginally. Therefore at the next instant the change in F is exactly $-\gamma$.

[2]The present value of the co-state however is zero when $t \to \infty$.

[3]If the optimal trajectory begins above the stationary locus for ψ then A^* is increasing for some time after $t = 0$ but eventually crosses $\dot{\psi} = 0$ in the first phase and A^* is decreasing up to T.

[4]If the duration of the first phase is not long enough then this may not happen, see for example ω' in Fig.1 and Fig.2.