Economic Research Institute Study Paper
ERI # 95-28

NONCONCAVE, NONMONOTONIC NETWORK EXTERNALITIES

By

Yi Nung Yang
Christopher B. Barrett

DEPARTMENT OF ECONOMICS
UTAH STATE UNIVERSITY
LOGAN, UTAH

July 1997

NONCONCAVE, NONMONOTONIC NETWORK EXTERNALITIES

Yi-Nung Yang
National Taiwan University
and
Christopher B. Barrett*
Utah State University

June 1997 revision

JEL codes: L11, L23
Keywords: critical mass, history-dependence, imperfect information, multiple equilibria, network externalities

Abstract: While the existing literature confines attention to monotonic and strictly (globally) concave network externalities, we make a case for considering nonconcave and nonmonotonic externality functions. This formulation offers a more intuitive depiction of the relationship between network size and adoption externalities, in part because it captures both the exclusivity and functionality values of network goods. Our approach also highlights the instability and history-dependence of network equilibria and the need for critical mass in network entry, unifying in one simple framework some of the most intriguing findings extant in the burgeoning literature on networks. .

* Barrett is corresponding author: phone (435) 797-2306, fax: (435) 797-2701, E-mail: cbarrett@b202.usu.edu. We thank Nicholas Economides, Chris Fawson, Neil Gandal, Terry Glover, William Neilson, Dawn Thilmany, Quinn Weninger and participants at the 1996 Western Economic Association International annual conference for helpful comments. This research was supported by the Utah Agricultural Experiment Station and approved as journal paper number 4864.

NONCONCAVE, NONMONOTONIC NETWORK EXTERNALITIES

I. Introduction

A rapidly growing literature considers the economics of industries based on products characterized by externalities in which the utility derived by consumers is a function of the number of consumers.[1] Such externalities might arise for any of a number of reasons: because the usefulness of the product depends directly on the size of the network (e.g., telephones, fax machines), or indirectly through the availability of complementary goods and services (often known as the "hardware-software paradigm") or of postpurchase services (e.g., for automobiles). Collectively, these phenomena are often termed "network externalities." These externalities generally cause the competitive market equilibrium network size to be socially suboptimal.

Network externalities are typically modeled as monotonic, strictly (globally) concave functions of consumer's expectations of network size. But other than for reasons of tractability, why should analysts assume concavity and monotonicity in network externalities? We believe such assumptions are often counterintuitive in this setting. Casual observation of paradigmatic networks like telephones, electronic mail, financial exchanges or computer software suggests the externality effects of a network are modest, potentially even decreasing, while the network is small, then increase quite rapidly once the network reaches some critical mass, and then gradually taper off again. At particularly large network sizes, externality effects may even turn negative due to overloading. That is, many network externality functions faced by a prospective consumer are unlikely to be strictly

[1] Especially noteworthy contributions include Katz and Shapiro [8; 9; 10], Farrell and Saloner [6; 7], Economides and Salop [4], Liebowitz and Margolis [11], Economides and Himmelberg [4], and the October 1996 issue of the *International Journal of Industrial Organization*.

concave, and perhaps not even monotonic. Perhaps this sounds like a quibbling point regarding the technical details, in particular the functional form, of a relatively recent class of microeconomic models. But these points are central to understanding the dynamics of adjustment to network equilibrium as well as the properties of the resulting equilibria, particularly under imperfect information or in the presence of liquidity constraints or bounded rationality. We do not claim all network externalities conform to the specification we introduce here, but we believe this simple technical refinement adds some interesting insights to a burgeoning literature on a subject of considerable importance to crucial international industries. We offer little in the way of truly new qualitative findings; ours is a methodological innovation that provides a less complex, more intuitive, unifying and tractable mechanism to reach conclusions generally scattered through the extant literature on networks.

II. Consumer Choice and Market Equilibrium

To make our points clearly, we work with the simple case of a single system under monopoly. This necessarily excludes the interesting issues of inter-firm coordination, competition and product compatibility, but focuses attention on the nature of the network externalities themselves.

In this section, we offer a brief and necessarily incomplete summary of the extant network externalities literature, which proceeds roughly as follows. The ith consumer's willingness to pay for the product, x, takes the form $r_i+v(x^e)$, where r_i is the consumer-specific myopic valuation of the product and $v(x^e)$ is the network externality value defined over expected network size. Consumers are assumed to value network externalities and form expectations over market size identically. The existing literature assumes $v' > 0$, $v'' < 0$, $v(0) = 0$ and $\lim_{x\to\infty} v'(x^e)=0$, i.e, concavity and monotonicity

in the network externality function. This form of network externality is depicted in the top panel of Figure 1.

The manner in which consumers form expectations about network size obviously influences market equilibrium.[2] Katz and Shapiro [8] employed — and most subsequent research has followed by using — a "fulfilled expectations" concept; consumers form rational expectations which must be satisfied in equilibrium. Even with a fulfilled expectations solution concept, multiple equilibria naturally emerge (one of which is no sales). But assuming non-instantaneous adjustment to shocks, including network introduction, one faces a dynamic problem of adjustment to equilibrium. This issue is generally overlooked in the existing network externalities literature, which implicitly presumes instantaneous adjustment. Furthermore, the dynamic problem of network size adjustment arises if there is imperfect information such that consumers' expectations are only "quasi-fulfilled", which we define as expectations that are fulfilled in the subsequent period, but not necessarily at an equilibrium value. Under such circumstances, imperfect information that might cause consumers to expect on a nonequilibrium network size can lead to a sequence of revised expectations that converge on a true equilibrium. Thus, consumers purchase the good at time t if and only if $r_i + v(x^e_t \mid \Phi_t) \geq p_t$ where p_t is the price charged and Φ_t is information available at time t, including x_{t-1}. As long as consumer expectations are not fulfilled, $\{x_t, p_t\}$ adjust to a dynamic equilibrium $\{x^*_t, p^*_t \mid p_{t-1}, \Phi_t\}$ conditional on the past price (p_{t-1}) at which the product is offered on the market and the information set that conditions consumers' expectations.

[2] If there exist sunk costs or economies of scale in production, the very existence of a good or service network depends on firms' expectations of consumer demand, which is in turn a function of individual consumers' expectations of others' demands. The (potentially heterogeneous) way in which agents form expectations is thus of extraordinary significance in industries characterized by network externalities.

Assume the firm can provide enough x to satisfy demand fully over the product life cycle.[3] The dynamic adjustment of the network size, i.e., the network growth function, can thus be captured in the differential equation $dx/dt = G(v(x^e_t),p_t)$. Assume G(.) initially increases in v(.), then decreases after some point K_2.[4] This is depicted in the middle panel of Figure 1. The sign of $\partial G(.)/\partial x^e_t=(\partial G/\partial v)v'$ depends on $\partial G/\partial v$ since $v'>0$ in the traditional model. Assume also that $(\partial^2 G/\partial v^2) < 0$ and $\partial^2 G(.)/\partial(x^e_t)^2 = (\partial G/\partial v)v''+ (\partial^2 G/\partial v^2)(v')^2 < 0$ in order to guarantee the convergence of the system. Assuming linear partial demand (r_i), and aggregating across a finite population of consumers, the concavity of the network externality and G(.) functions generate a concave network growth function with an equilibrium $x_t^* > K_2$, as shown in the bottom panel of Figure 1. This is the conventional result, with details commonly implicitly assumed brought into the open.

The monopoly firm chooses an optimal pricing strategy to maximize the discounted present value of the profit stream, taking into account the impact of the different network growth paths associated with different prices. Thus the firm's objective is

$$\text{Max}_{p_t} \int^T_0 e^{-rt} (p_t - c)y_t \, dt \tag{1}$$
$$\text{s.t.} y_t=G(v(x^e_t),p_t) \tag{2}$$
$$dx^e_t/dt = y_t \tag{3}$$
$$x_T= \int^T_0 y_t \, dt \geq 0 \tag{4}$$
$$x^e_0=0 \tag{5}$$

where r is the discount rate, c is the constant marginal cost of production, and y_t is sales flow. Note that under the earlier assumptions regarding G(.) the monopolist need not worry about scale of entry

[3] The possibility of overlapping technology generations, and thus of upgrade timing choices under both concave and nonconcave externalities, is the subject of separate research by the senior author.

[4] Given the assumption of a monotonic, concave network externality function, $V(x_t^e)$, the network will converge to a finite equilibrium only if G(.) is at some point decreasing in $V(x_t^e)$. This point has not been recognized in the literature to date. A proof is available from the authors by request.

since the growth path automatically leads to a stable, positive equilibrium network size. Defining $\phi_t = e^{rt}\lambda_t$, where λ_t is the costate variable, the current value Hamiltonian can be written as $H=(p-c+\phi_t)G(.)$. The maximum principle suggests that the necessary conditions for a solution to the current value Hamiltonian are

$$(p_t - c + \phi_t)\ \frac{\partial G}{\partial p_t} + G = 0 \tag{6}$$

$$\frac{d\phi_t}{dt} = r\phi_t - (p_t - c + \phi_t)\ \frac{\partial G}{\partial v}\frac{\partial v}{\partial x_t^e} \tag{7}$$

$$\frac{dx_t}{dt} = y_t = G(v(x_t), p_t) \tag{8}$$

Furthermore, the transversality condition tells us that $\phi_t \geq 0$, $\phi_T\ x(T) = 0$.

Assuming an interior solution[5], the necessary conditions for maximization of the monopolist's profits yield the marginal revenue equals marginal cost condition

$$p^*[1+1/\varepsilon_t]+\phi_t = c \tag{9}$$

where ε_t is the price elasticity of demand. Network externality effects, captured in ϕ_t, have the effect of reducing the monopolist's optimal price, p_t^*. So if the network externality function is strictly concave, then λ_t is decreasing in network size, ϕ_t is decreasing in time, and the monopolist's markup is increasing over time (holding elasticity constant). This point has been largely overlooked in the literature to date although it has interesting implications for the regulation of natural monopolies in network goods and services (e.g., electricity, telecommunications).

III. Nonconcave Network Externalities

[5] The second order condition requires $\partial^2 H/\partial p_t^2 = (p_t - c+\phi_t)\ (\partial^2 G/\partial p_t^2) + 2(\partial G/\partial p_t) < 0$

Now consider what happens if network externalities are not globally concave and monotonic. For example, assume instead that $v(x_t^e)$ follows the critical depensation form from bioeconomics [3] as shown in the top panel of Figure 2. Formally, assume this nonconcave, nonmonotonic network externality function exhibits $v'<0$ when $x^e < K$ and $v'>0$ when $x^e > K$, and $v'' > 0$ if $x^e < x^*_1$ and $v'' < 0$ if $x^e > x^*_1$, where x^*_1 is the critical mass. This network externality is depicted in the top panel of Figure 2.

Why might the network size externality take such a shape? Recognize that network size affects both the functionality of a good and its exclusivity. To date the networks literature has assumed consumers value only the former. But casual observation of purchasing patterns and consumer psychology when network goods are introduced suggests that initial purchasers place considerable positive value on the exclusivity of their possessions. New technologies, network technologies not excepted, are often status goods when they are held by only a few. This exclusivity value necessarily decreases in network size, going to zero once enough people own the good. The functionality value associated with network size remains very low initially then increases rapidly once the network achieves some minimum effective size. This functionality value may level off at some point once the network becomes saturated. It might even decrease if the network becomes overloaded.[6] We implicitly model the exclusivity value as dominating the functionality value up to some positive mass K. After that point, $v(x_t^e)$ is increasing due to increasing functionality value. Meanwhile, any externalities related to the exclusivity of the network good are soon exhausted.

[6] We do not model the case of overloaded networks, wherein the externality v' turns negative once again. This is nonetheless a relatively straightforward extension.

With that sort of a nonconcave, nonmonotonic network externality, and maintaining the same functional form used earlier for G(.) (reproduced as the middle panel of Figure 2), the network growth function takes on a markedly different shape, as depicted in the bottom panel of Figure 2. At least three distinct equilibria now emerge with the nonconcave network externality. First, x*=0, network extinction, is a stable equilibrium, as one can prove by discretizing and differentiating equations (3) and (4).

$$\begin{gathered} dx_t^e/dt=G(v(x_t^e),p_t) \\ x_{t+1}^e - x_t^e = G(v(x_t^e),p_t) \\ x_{t+1}^e = x_t^e + G(v(x_t^e),p_t) \\ \partial x_{t+1}^e / \partial x_t^e = 1+(\partial G/\partial v)\, \partial v/\partial x_t^e \end{gathered} \tag{10}$$

Since $\partial G/\partial v > 0$ and $\partial v/\partial x_t^e < 0$ in the neighborhood of v(0) by earlier assumptions, $\partial x_{t+1}^e/\partial x_t^e < 1$ and $x^* = 0$ is a stable equilibrium. A stable, positive equilibrium occurs at $x_2^* > K_2$, where $\partial G/\partial v<0$ and $v' > 0$. These are the finite, stable equilibria of the traditional literature on network externalities [8,9]. The nonconcave network externality function, however, generates an unstable, positive equilibrium network of size $x_1^* \in (K_1,K_2)$ because

$$\partial x^e_{t+1} / \partial x^e \,|_{x^e = x^*_1} = 1 + (\partial G/\partial v)v' > 1 \tag{11}$$

(evaluated around the point $x^e = x_1^*$). Additional unstable positive equilibria could result at superlarge sizes if v' again turns negative due to network congestion and overloading externalities. Those are not depicted in Figure 2.

The shape of the network externality function thus fundamentally changes the dynamics of network growth. This simple refinement thereby adds a number of useful characteristics to a basic network model that otherwise require more complex and less intuitive modeling. If the product is introduced with a network size less than x_1*, the critical mass associated with the inflection point in

$v(x_t^e)$, it regresses irreversibly toward extinction, not toward the stable, positive-output equilibrium size of x_2*. It is also evident that $d\phi/dt$ is positive over the range $[0,K^+]$, where $d\phi/dt=r\phi$ (equivalently, $d\lambda/dt=0$) at $x^e = K$ ($K^+ > K$) according to (7). In words, the monopolist's mark-up will now be time-varying, temporarily decreasing to some intermediate network size — as is commonly the case with introductory offers and, in the limit, distribution of network products as freeware — then increasing once the monopolist has cleared the critical mass threshold.

Under the assumption of fulfilled expectations and instantaneous adjustment to equilibrium, the unstable, intermediate equilibrium would not result for the simple reason that it is not a subgame perfect equilibrium. A monopolist with perfect foresight would choose the stable, positive equilibrium network size, x_2* — or at least an initial network size that would converge toward x_2* — since this is the only equilibrium offering stable, positive profits. Thus the unstable equilibrium is a set of measure zero if there is perfect information and the capacity to act on that information.

Under quasi-fulfilled expectations due to imperfect information, however, or in the case where the firm faces liquidity, informational or bounded rationality constraints, then the intermediate, unstable equilibrium becomes a point of real interest, not a mere abstraction. If consumers significantly adjust expectations over time, if the monopolist significantly misjudges consumer expectations — whether due to misinformation or error of judgement — or if the monopolist is unable to finance potential termporary losses necessary to select a one period flow sufficient to hurdle the nonconcavity and put himself on a saddle path toward x_2*, then the intermediate, unstable equilibrium becomes a real possibility. Given the likelihood of these complications, we think this a particularly interesting possibility of considerable practical interest to the analysis of network industries.

Nonconcave, nonmonotonic externalities make several crucial, existing insights more transparent, tractable and integrated and adds some extra insights as well. First, other researchers, notably Besen and Farrell [1], have described network markets as "tippy," i.e., prone to instability. But they have had to approach this point through far more complex analytics. By generating an inflection point in the network externality function by combining exclusivity and functionality values, our nonconcave network externality function generates such tippiness directly in its unstable equilibria. Momentarily extending the model to potential oligopoly, if competing networks emerge, ultimately settling into equilibria of different sizes, those who joined what turns out to be the smaller network may switch to the other network (if the marginal benefit of the larger network exceed the switching costs), thereby perturbing the smaller network from its unstable equilibrium and toward the stable equilibrium of extinction.[7] Cases such as Supercalc spreadsheet software, Wang computers, and Betamax videocassette recorders spring immediately to mind as examples. More innovatively, even in the monopolistic case we have modeled, a negative shock to demand can precipitate the demise of a network poised at or near an unstable, intermediate-sized equilibrium. The ill-fated New Orleans rice futures market offers an excellent example of this phenomenon, which the existing literature does not capture.

Second, the existence of unstable intermediate equilibria reinforces claims that scale of entry matters enormously to firms providing products characterized by network externalities, irrespective of production costs. Economides and Himmelberg [4] raise this issue of critical mass, showing (again in the oligopolistic case) that only the bigger of two networks of different size is Pareto dominant and

[7] Church and Gandal [2] generate such effects using concave (CES) indirect utility functions in a competition between incompatible software packages.

stable. If one adopts a nonconcave network externality function, particularly the specification presented above, such critical mass emerges naturally[8]. Insufficient initial network size (below x^*_1 in Figure 3) leads to depressed product demand as consumers adjust downward their expectations of network size, and hence the value added from adoption externalities. Firms advertise, distribute a large number of pre-release ("beta") versions of a product, tolerate product piracy or employ other marketing tactics to ensure sufficient scale of entry to prevent regress toward extinction.

Third and following up on the second point, nonconcave network externalities add another dimension to the gains from market dominance. Not only might the firm with the largest network in an industry be able to extract greater profit due to the added value consumers derive from its larger network or from lower unit costs in the presence of scale economies, but it may have leapt over the unstable intermediate equilibrium to the larger, stable one. A higher probability of survival, not just short-term profits, accrue to larger networks if adoption externalities are nonconcave. This issue of stability, and thus sustainability, sheds a bit more light on both the fierce battle over product standards in races to introduce new technologies (e.g., videocassette recorders, digital audio format), on the need for late-comers to maximize compatibility with the existing market leader (e.g., Borland's mimicking of Lotus when introducing Quattro), and on the incentives for firms to overestimate adoption rates for their own technologies (as occured in the battle between IBM and Microsoft over the Windows and PS/2 operating systems). It also helps explain why firms might be more receptive to letting others copy their technologies, as freeware or shareware (e.g., Mosaic and Netscape in

[8] Critical mass also arises naturally in Economides and Himmelberg [4] through their assumption of a multiplicative functional form for consumer demand with globally strictly concave network externalities. As a consequence, their fulfilled-expectations demand function is not monotonic.

Internet software) or be lenient with software pirates in early efforts to establish a large, sustainable network size [12].

Fourth, several commentators on network externalities refer to the history-dependence of network size [1]. Hysteresis emerges quite naturally from modelling the network externality effect as following the nonconcave form we have laid out. Recall that the slope of the growth function, $G(v(x^e_t),p_t)$, depends fundamentally on x^e_t. Thus the stable equilibrium (x^*_0 or x^*_2) toward which the network converges is conditional on x^e_t. The disequilibrium dynamics of networks exhibiting nonconcave externalities turn on the product's (equilibrium and disequilibrium) history.

Fifth and finally, nonconcave network externalities generate a non-monotonic time path of monopoly markups. Since λ_t increases over some range of x_t, ϕ_t increases in time over part of the product's life. Our nonconcave structure to network externalities leads to a partial cycling of monopoly markups: they decrease from product introduction through some intermediate network size, then begin growing again. Casual observation suggests this is frequently the path of markups (or pricing in the case of constant or decreasing costs).

IV. Conclusions

One's assumptions about the nature of network externalities matter to the resulting equilibria and to the disequilibrium dynamics of network size adjustment. While the existing literature confines attention to strictly concave and montonic network externality functions, we make a case in this paper for considering also nonconcave and nonmonotonic forms, in particular the critical depensation form from bioeconomics [3]. This follows from the observation that network size affects not just the functionality of the good but also its exclusivity, which some consumers value. This allows us to be

more specific — relative to the existing literature — about the relationship between network effects and the size of the network and to capture in a simple model some of the most intriguing insights of this burgeoning literature.

What insights does this simple technical refinement offer? Aside from a more intuitive depiction of the relation between network size and adoption externalities, nonconcavity provides explicitly for unstable equilibrium, critical mass in network entry, the importance of market dominance, and history-dependence, all important but to-date disconnected findings. This approach thus helps unify core findings on network externalities. It also yields interesting insights on the non-monotonic time path of monopoly mark-up pricing for network products, a point thus far absent in the literature.

REFERENCES

1. Besen, S.M. and J. Farrell, "Choosing How to Compete: Strategies and Tactics in Standardization." *Journal of Economic Perspectives*, 1994, 8, 117-131.

2. Church, J. and N. Gandal, "Network Effects, Software Provision, and Standardization." *Journal of Industrial Economics*, 1992, 40, 85-104.

3. Clark, C. W., *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*, second edition (New York: John Wiley & Sons, 1990).

4. Economides, N. and C. Himmelberg, "Critical Mass and Network Evolution in Telecommunication." in G. Brock, ed., *Toward A Competitive Telecommunications Industry: Selected Papers From The 1994 Telecommunications Policy Research Conference*, 1995.

5. Economides, N. and S. Salop, "Competition and Integration Amond Complements, and Network Market Structure." *Journal of Industrial Economics*, 1992, 40, 105-123.

6. Farrell, J. and G. Saloner, "Standardization, Compatibility, and Innovation." *Rand Journal of Economics*, 1985, 16, 70-83.

7. Farrell, J. and G. Saloner, "Installed Base and Comptability: Innovation, Product Preannouncements, and Predation." *American Economic Review*, 1986, 76, 940-955.

8. Katz, M. L. and C. Shapiro, "Network Externalities, Competition, and Compatibility." *American Economic Review*, 1985, 75, 424-440.

9. Katz, M. L. and C. Shapiro, "Technology Adoption in the Presence of Network Externalities." *Journal of Political Economy*, 1986, 94, 822-841.

10. Katz, M. L. and C. Shapiro, "Systems Competition and Network Effects". *Journal of Economic Perspectives*, 1994, 8, 93-115.

11. Liebowitz, S. J. and S.E. Margolis, "Network Externality: An Uncommon Tragedy." *Journal of Economic Perspectives*, 1994, 8, 133-150.

12. Takeyama, L. N., "The Welfare Implications of Unauthorized Reproduction Of Intellectual Property In The Presence of Demand Network Externalities." *Journal of Industrial Economics*, 1994, 42, 155-165.

Figure 1: Network Growth With Concave Network Externalities

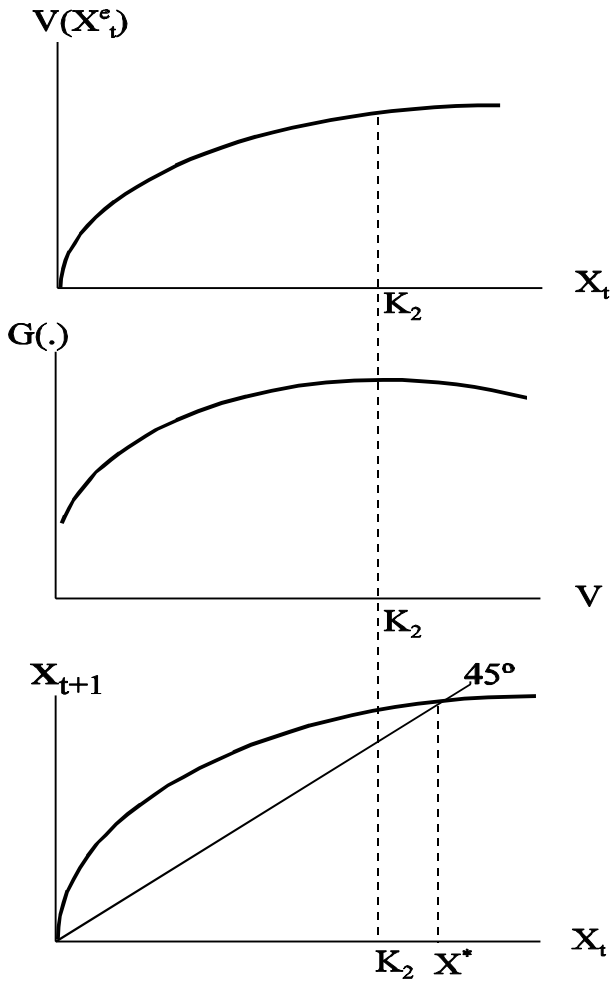

Figure 2: Network Growth With Nonconcave Network Externalities