Forthcoming in *Toward a Competitive Telecommunications Industry: Selected Papers from the 1994 Telecommunications Policy Research Conference*, Gerard Brock (ed.), 1995.

Critical Mass and Network Evolution in Telecommunications

by

Nicholas Economides and Charles Himmelberg

December 1994

Abstract

Network goods exhibit positive consumption externalities, commonly known as "network externalities." In markets, this fact can give rise to the existence of a *critical mass point*, that is, a minimum network size that can be sustained in equilibrium, given the cost and market structure of the industry. In this paper, we describe the conditions under which a critical mass point exists for a network good. We also characterize the existence of critical mass points under various market structures. Surprisingly, neither existence nor the size of the minimum feasible network depends on market structure. Thus, even though a monopolist enjoys an additional degree of freedom through its influence over expectations, and even though monopolistic and oligopolistic markets will in general provide a smaller sized network than perfect competition, the critical mass point is nonetheless the same. We extend these results by making the model dynamic and by generalizing it to allow durable goods. Introducing network externalities to a dynamic model of market growth increases the speed at which market demand grows in the presence of a downward time trend for industry marginal cost. We use this prediction to calibrate the model and obtain estimates of the parameter measuring a consumer's valuation of the installed base (i.e., the network effect) using aggregate time series data on prices and quantities in the US fax market.

** Stern School of Business, New York University, New York, NY 10012-1126, (212) 998-0864 and (212) 998-0870, respectively. Fax: (212) 995-4812. Email: neconomi@stern.nyu.edu, and chimmelb@stern.nyu.edu.

Critical Mass and Network Evolution in Telecommunications

1. Introduction

Many consumption goods, most notably in the telecommunications industry, exhibit *network externalities*. That is, the value of the good to the consumer depends on the number of consumers purchasing the same (or a similar) good. Many other consumption goods, such as computer software, display network externalities, though usually in less-obvious ways. All such goods have in common the feature that an increase in expected sales of network services creates positive expected benefits for every consumer of the good.

In this paper, we show that consumption goods with network externalities are often characterized by the existence of a *critical mass point.* That is, an equilibrium market for the good does not exist unless the installed base is greater that a minimum level. We show that this is a general feature of goods that exhibit network externalities and will be observed in a variety of market structures. Nevertheless, the presence of positive network externalities and positive and significant critical mass have significant impact on the analysis of conduct, structure, and performance of network industries. We analyze single period as well as multi-period (dynamic) effects of network externalities both for perishable and durable goods.[1]

We illustrate these ideas using the U.S. market for facsimile machines as an example. This example illustrates the potentially large effects of network externalities on the growth behavior of markets and outlines a strategy for estimating the effect of externalities using such models. Specifically, we focus on the growth of the market following the introduction of the (industry accepted) D3 standard in the late 70s. During the early 80s, as fax machines conforming to this standard were just beginning to ship, our data indicate that fax machines were selling at an average price of more than $2,000. This high average price evidently exceeded the reservation price of most consumers, because by 1983 the estimated installed base was still less

[1] Many of the results of this paper are based on Economides and Himmelberg (1994).

than one million machines. But by 1983, the average selling price had begun to decline, precipitously. Between 1982 and 1985, average prices fell by a factor of four to about $500, and by 1987, had fallen still further to about $250.

This dramatic price drop was driven by dramatic reductions in the price of microelectronic components used in the production of facsimile machines, and it sparked an explosion in demand. Prior to this point in time, most of the demand had come from "early adopters," which had been sufficient to keep the market growing during the first half of the 80s at a rate of about 20 percent per year. But in 1986, the demand for fax machines began to accelerate dramatically; in 1987 it exploded to more than double the previous year, and in 1988 demand more than doubled again. Demand for fax machines remained robust throughout the rest of the decade, and by 1991 the installed base had grown to more than 10 million machines.

While the dramatic drop in prices clearly played a major role in the growth of the fax market, we argue that the particularly explosive growth during the mid-to-late 80s was fueled by both realized and anticipated increases in the size of the installed base. In the model below we derive the aggregate demand function from a discrete choice model of demand at the level of the individual consumer. This "bottom-up" approach allows us to infer the effect of network externalities on consumer purchasing decisions. The model assumes that the *net* value of a fax machine to an individual consumer depends on that consumer's income level, the size of the installed base, the price paid for the machine, and a random element the captures idiosyncratic tastes. We use this model to predict the rate at which the installed base should have grown in the absence of network externalities, given the path of prices and the distribution of consumer income. We then compare the theoretical rate of market growth predicted by the model with the empirical rate of market growth in the data. On the basis of the discrepancy between these two, we back out the implied parameters of consumer utility function that measure the value of the installed base.

In the next section, we begin our discussion by introducing some basic conceptual framework for thinking about network externalities with consumer goods, and we focus on the nature of communication networks in particular. In the following section we introduce the intuition for critical mass points as well as the intuition for the formal conditions under which network goods would display such phenomena. We show that the critical mass of a network is (surprisingly) independent of market structure, and that market equilibria are (unsurprisingly) generically inefficient relative to the allocation that would maximize total welfare. In sections 3 and 4, we show that our static framework generalizes to a dynamic setting, and for the benefit of our empirical exercise, we further modify the framework to accommodate the durability of fax machines.

2. Sources of Network Externalities

The key reason for the existence of network externalities is the complementarity between the components of a network. Depending on the network, the externality may be direct or indirect. Networks where services AB and BA are distinct are named *two-way networks* in Economides and White (1993). Two-way networks include railroad, road, and many telecommunications networks, such as the facsimile network that will analyze in later sections. In typical two-way networks, customers are identified with components and the externality is direct. Consider, for example, the local telephone network of Figure 1. In the n-component network of Figure 1, there are $n(n - 1)$ potential goods. An additional customer provides *direct externalities* to all other customers in the network by adding $2n$ potential new goods through the provision of a complementary link

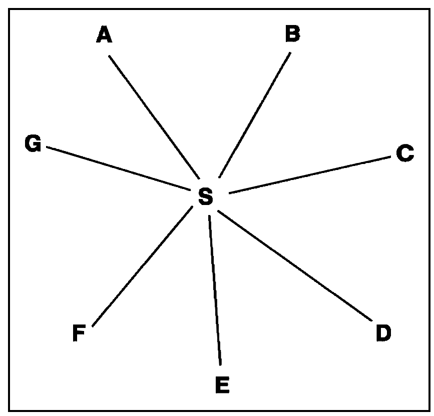
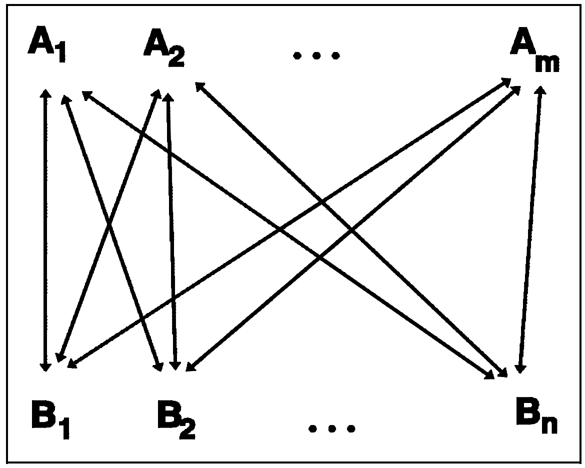


Figure 1: A local telephone network in a star topology.

(say GS) to the existing links.[2] For example, in a simple fax network with, say, 100 nodes, there are 9900 distinct "goods" available (99 per node), and the addition of the 101st node creates an additional 202 goods (distinct, one-way fax transmissions).



<u>Figure 2</u>: A pair of vertically-related markets.

When one of AB or BA is unfeasible, or does not make economic sense, or when there is no sense of direction in the network so that AB and BA are identical, then the network is called a *one-way network.* In a typical one-way network, there are two types of components, and composite goods are formed only by combining a component of each type, and customers are often not identified with components but instead demand composite goods. For example, Automatic Teller Machine (ATM) services, broadcasting television (over-the-air and cable), electricity networks, retail dealer networks, and paging are one-way networks.

In typical one-way networks, *the externality is only indirect*. When there are m varieties of component A and n varieties of component B as in Figure 2 (and all A-type goods are compatible with all B-type), there are mn potential composite goods. An extra customer yields indirect externalities to other customers, by increasing the demand for components of types A

[2] This property of two-way networks was pointed out in telecommunications networks by Rohlfs (1974) in a very early paper on network externalities. See also Oren and Smith (1981) and Hayashi (1992).

and B and thereby (because of the presence of economies of scale) potentially increasing the number of varieties of each component that are available in the market.[3]

2. Critical Mass

For normal goods that do not exhibit network externalities, demand slopes downward; as price decreases, more of the good is demanded. Conversely higher levels of consumption are associated with lower prices. This fundamental relationship may fail in goods with network externalities. For these goods, the willingness to pay for the last unit increases as the number expected to be sold increases. If expected sales equal actual sales, the willingness to pay for the last unit *may* increase with the number of units sold. Thus, for goods with network externalities, the (fulfilled expectations) demand-price schedule may not slope downward everywhere. In such markets, as costs decrease we may observe discontinuous expansions in sales rather than the smooth expansion along a downward slopping demand curve. In particular, we may observe a discontinuous start of the network: as costs decrease, the network starts with a significant market coverage (say 10% of the market) rather than starting with 0.1% coverage.

Critical mass is defined as the minimal non-zero equilibrium size (market coverage) n^0 of a network good or service (for any price). We will argue that, for many network goods, the critical mass is of significant size, and therefore for these goods small market coverage will never be observed -- either their market does not exist or it has significant coverage.

The concept of critical mass formalizes the "chicken and the egg" paradox that logically arises in such markets, namely: many consumers are not interested in purchasing the good because the installed base is too small, and the installed base is too small because an insufficiently small number of consumers have purchased the good. Thus, consumers'

[3] In many industry structures, the addition of new varieties is concurrent with an intensification of competition; in these cases, consumers have the added benefit of price decreases as the number of varieties increases.

expectations of no network good provision may be fulfilled. However, for a range of costs, expectations of positive level(s) of sales of the network good are also fulfilled. Often, there are multiple fulfilled expectations equilibria. Consumers and producers can coordinate to reach any one of them. We will assume that they will reach the equilibrium of the largest network size. Thus, when more than one network size is supported by the same price, we select as the equilibrium the highest network size supported by that price; this network size Pareto dominates the other network sizes supported by the same price.[4]

2.1 Perfect Competition

The essential features of our model are derived from assumptions about the utility that consumers receive from owning the network good. For simplicity, assume that utility that consumers receive from owning the network good is proportional to their income, y, and further assume that their willingness to pay for the good is given by the function $u(y, n) = y^{\gamma}(k + dn^{\alpha})$, where n is the market coverage of the network (installed base), $n \in [0, 1]$, α takes values between zero and one, and d is a parameter with a value greater than zero. Without loss of generality, we assume $\gamma = 1$. The constant term k is the innate value of the good to the consumer when the size of the installed base is zero. This functional form assumption implies that utility is increasing with the size of the installed base, but that its marginal value is decreasing if $\alpha < 1$. This is a rough approximation intended to capture the intuition that individuals with higher incomes would tend to make more use of fax machines and therefore place a higher value on both the machine itself and the size of the installed base.[5]

---

[4] Our analysis shows that the equilibrium is also stable.

[5] This specification of the utility function can be readily generalized to accommodate an arbitrarily large number of consumer characteristics associated the utility derived from faxes. We chose income because data on the distribution of incomes are readily available and because income is probably the best single measure of willingness to pay. Using a univariate also simplifies the exposition substantially.

From these assumptions about individual utility, we derive the aggregate demand by counting the number of people willing to purchase the good given the market price p and the size of the installed base n^e, i.e., the number of people with incomes sufficiently high so that $u(y, n) \geq p$, where p is the price of the good. Thus, aggregate demand takes the general form $n = f(n^e, p)$. For a given level of the installed base, n^e, this demand function is downward sloping in price. As the expected size of the network n^e increases, this demand curve shifts up.

We can invert the demand curve above to yield the price that the marginal consumer is willing to pay given the installed base and the number of people who demand the good, that is, $p = p(n, n^e)$.[6] Thus for a given level of the installed base, n^e, price is a function of n as indicated by the downward sloping curves plotted in Figure 3.

Up to this point in the discussion we have allowed the size of the installed base to differ from the size of the level of network. In equilibrium, this would clearly be impossible, since the level of network demand obviously determines the size of the installed base of the market clears. Therefore, we now impose the equilibrium condition that $n = n^e$, and we refer the resulting demand curve as the *fulfilled expectations demand,* $p(n, n)$. Figure 3 shows the construction of a typical fulfilled expectations demand. The curves $p(n, n_1^e)$ and $p(n, n_2^e)$ show the willingness to pay, given different sizes of the installed base that consumers expect to emerge in equilibrium, where $n_2^e > n_1^e$. The point labeled E_1 on the first curve represents the point at which n equals n_1^e, and analogously, E_2 on the second curve represents the point at which n equals n_2^e. The locus of all such points traces out the fulfilled expectations demand curve. Observe that *the fulfilled expectations demand $p(n, n)$ is not monotonic*. Also note that, in addition to the prices

[6] The willingness to pay function can be derived as follows. We assume that total demand is given by the total number of consumers for whom $u(y, n) > p$, i.e., the number of consumers with incomes satisfying $y > y^*$, where $y^* = p/(k + dn^\alpha)$. Therefore demand is given by $n^d = 1 - G(p/(k + dn^\alpha))$, where G is the distribution of income. The willingness to pay function is derived by solving for p, so that $p(n, n^e) = (k + dn^\alpha)G^{-1}(1 - n)$, and the fulfilled expectations willingness to pay function is $p(n, n) = h(n)G^{-1}(1 - n)$.

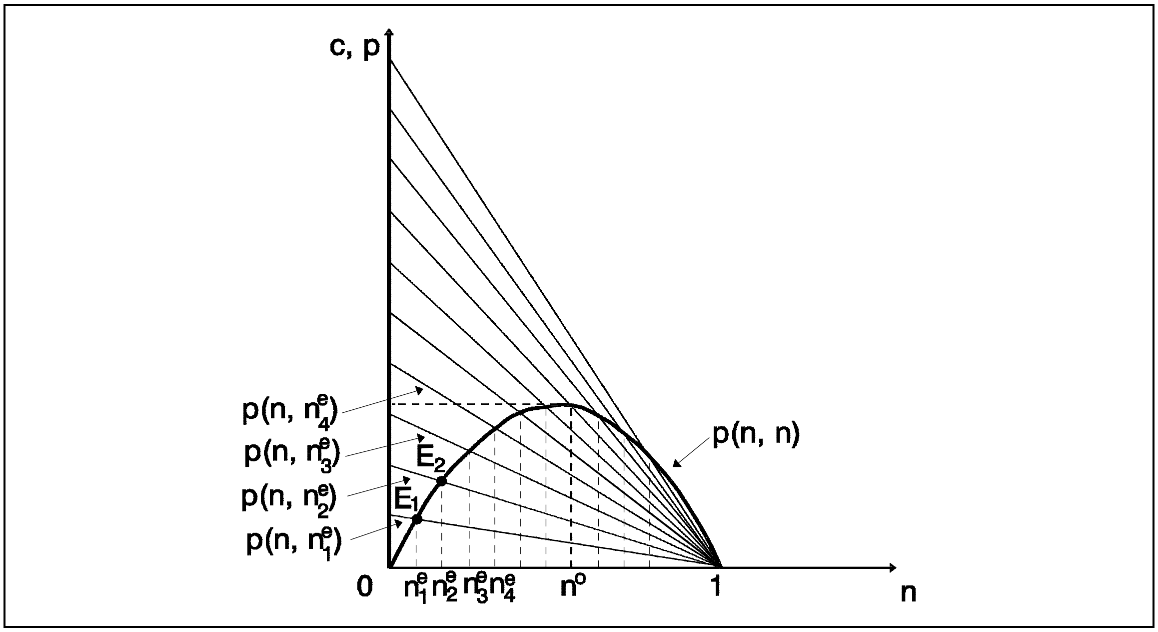


Figure 3: Construction of the fulfilled expectations demand.

indicated by the inverted-U shaped curve in Figure 3, the fulfilled expectations demand curve p(n, n) also includes the entire vertical axis at zero, which is drawn thicker on purpose. This is because at any marginal cost $c > k$ a network of zero size is a fulfilled expectations equilibrium, and Figure 3 is drawn for the special case when $k = 0$.

For $n > 0$, the fulfilled expectations demand is single-peaked (quasi-concave) for a fairly general set of conditions. We can show that $\lim_{n\to 1} p(n, n) = 0$, so that p(n, n) is decreasing for large n. This guarantees that the market does not explode towards infinite output. Given single-peakedness, the fulfilled expectations demand can either decrease everywhere, or have an increasing part for small n, as in Figure 3. If p(n, n) decreases for all n (the case of weak network externalities), it exhibits no qualitative difference to an ordinary demand curve. The interesting case arises when p(n, n) increases for small n (the case of strong network externalities).

Figure 4 shows the fulfilled expectations demand for strong and weak network externalities. Focusing on the case of strong network externalities pictured on the left side,

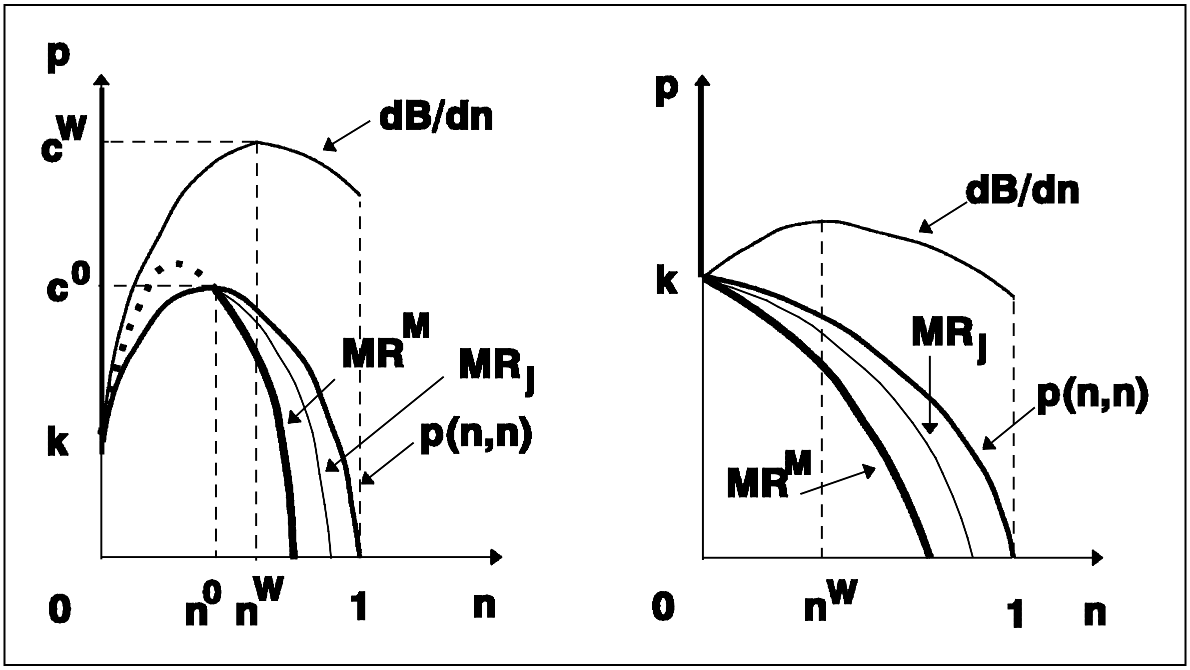


Figure 4: **The fulfilled expectations demand with strong and weak network externalities.**

consider perfect competition with constant marginal cost c. In equilibrium, price equals marginal cost, i.e.,

$$p(n, n) = c.$$

Let c^0 denote the peak of $p(n, n)$. For $c > c^0$, the only equilibrium is of zero size. For $c^0 > c > k$, there are two other equilibria, besides the zero one, at the intersections of the horizontal at c with $p(n, n)$. The lower of the two is unstable, and the higher one is stable. For $c < k$, there is only one equilibrium, and it is positive and stable. When more than one network size is supported by the same price, we select as the equilibrium the highest network size supported by that price. Thus, the equilibrium we select is always Pareto dominant and stable. In the case of strong network externalities, for $c > c^0$, the equilibrium is of zero size; at $c = c^0$, the network starts at size $n^0 > 0$; and, for $c < c^0$, the network follows the outer part of $p(n, n)$ with size $n > n^0$. Clearly, the market on the left panel of Figure 4 exhibits a positive and significant critical mass of size n^0. Thus, non-zero networks of smaller size than the critical mass, $n < n^0$ and $n \neq 0$, will not be observed at any prevailing marginal cost or price.

Under what conditions do network services exhibit critical mass? The crucial requirement is that the fulfilled expectations demand is increasing for small n. Economides and Himmelberg (1994) show that the fulfilled expectations demand is increasing for small n if either one of the three following conditions hold:

1) The utility of every consumer in a network of zero size is zero;

2) There are immediate and large external benefits to network expansion for very small networks;

3) There is a significant density of high-willingness-to-pay consumers who are just indifferent on joining a network of approximately zero size.

The first condition is straightforward and applies directly to all two-way networks, where network goods have no value if there are no other participants. The typical example is a telephone or fax network. These networks exhibit critical mass, i.e., they start with a significant market coverage. The second condition describes goods that may have some intrinsic value in zero-size networks, but their value increases dramatically as sales expand. A good example of this may be a specialized computer program which relies on support mainly from other users.[7] The addition of even few users can increase significantly technical support and the value of the product. Another example is a specialized newsgroup on internet. The third condition describes goods that may have some intrinsic value in zero-size networks and their value does not increase dramatically as sales increase, but have very widespread appeal. A good example of this may a computer software with large sales, but low externality from each sale. Each extra copy sold of a word processing program, such WordPerfect, creates a small externality; however, once its

[7] Some freeware and shareware computer programs fall into this category. Support is provided essentially by other users through discussions on bulletin board services (BBSs).

sales become very large, secretaries get trained in WordPerfect, thus creating a significant externality.

2.2 Welfare Maximization

An social welfare maximizing planner can fully internalize the externality. Thus, we expect that the planner will provide a larger network size. The planner maximizes the fulfilled expectations net benefit of a network of size n,

$$W(n, n) = B(n, n) - C(n) = \int_0^n (p(q, n) - c)dq,$$

where $B(n, n) = \int_0^n p(q, n)dq$ is the gross benefit of the network. The optimal choice is defined by the network size that makes the marginal net benefit equal to zero,

$$dW/dn = dB(n, n)/dn - c = p(n, n) + \int_0^n p_2\, dq - c = 0,$$

where $p_2 > 0$ denotes the derivative of the gross benefit function B with respect to its second argument, i.e., with respect to expected sales. Thus, the marginal gross benefit is higher than the fulfilled expectations willingness to pay, $dB(n, n)/dn = p(n, n) + \int_0^n p_2\, dq > p(n, n)$. See Figure 4. It follows that *a planner will start the network for marginal costs, $c \in [c^w, c^0]$, for which there would be no network under perfect competition, and will always support a larger network than perfect competition.*

The wedge between price and gross marginal benefit, ignored by perfect competition, but taken into account by the planner, clearly implies that perfect competition is inefficient. Finally, as seen on the right panel of Figure 4, *the welfare maximizing solution may exhibit a positive and significant critical mass, even when there is no positive critical mass under perfect competition.*

2.3 Monopoly

A monopolist may or may not be influence expectations of consumers. If the monopolist does not influence expectations, it is clear that it will produce less than perfect competition and have greater inefficiency. On the other hand, there is hope that a monopolist who can influence expectations will support a larger network than perfect competition thereby resulting in higher total surplus. However, we will show that a monopolist who is unable to price discriminate will always support a network that is smaller than perfect competition and results in lower total surplus. The reason for this is simple. The monopolist has two opposite incentives in setting network size. It would like to increase the network size to increase the surplus it can appropriate. On the other hand, the monopolist wants to reduce quantity below the competitive level so that it can increase price. We show that the second incentive is dominant.

The monopolist's profits,

$$\Pi^M(n, n) = R^M(n) - C(n) = n(p(n, n) - c),$$

are maximized when

$$d\Pi^M/dn = MR^M - MC = p(n, n) + ndp/dn - c = 0.$$

The marginal revenue curve is shown in Figure 4 as a bold line for $n > n^0$, and as a dotted line for smaller n.

The monopolist will operate only when price is above marginal cost. This implies $dp/dn < 0$, and therefore $MR^M = p(n, n) + ndp/dn < p$. Thus, the monopolist will only operate on the downward slopping part of the fulfilled expectations demand. As for any downward slopping demand, in this portion, marginal revenue is below price and the quantity produced by the monopolist falls below that of perfect competition. Therefore *the monopolist starts the network service at the same cost as perfect competition and has the same critical mass.* For all smaller marginal costs, the monopolist produces less than perfect competition and charges a lower price.

Despite his influence on expectations, *a monopolist supports a network which is smaller and more inefficient than perfect competition* from a social welfare point of view. Therefore *the existence of network externalities cannot be claimed as a reason in favor of a monopoly market structure*.[8]

2.4 <u>Oligopoly</u>

Oligopolists may produce network goods that are all compatible to each other, or some firms may produce goods that are incompatible with some subset of goods of other producers. In this paper we will consider only compatible goods oligopoly. This is not because of lack of theoretical interest in the incompatible goods case (see Economides (1994)), but we want to focus our attention to the application to the fax market where there is no competing technical standard. Suppose that all firms produce compatible and identical goods. Consider Cournot oligopoly among them. Assume that each firm is able to influence the expectation of consumers only about its own quantity of production. Then it is easy to see that this oligopoly will result in an outcome that lies between perfect competition and monopoly.

Firm j's profits in an k-firm oligopoly,

$$\Pi_j = n_j[p(\Sigma_{i=1}^k n_i, n_j+\Sigma_{i\neq j}^k n_i^e) - c],$$

are maximized when

$$d\Pi_j/n_j = MR_j - MC = p(\Sigma_{i=1}^k n_i, n_j+\Sigma_{i\neq j}^k n_i^e) + n_j dp/dn_j - c = 0.$$

It is easy to show that the resulting fulfilled expectations equilibrium is symmetric and $n_i^e = n_i = n/k$, for all i, where $n = \Sigma_{i=1}^k n_i$, so that the equilibrium is characterized by

[8] Of course, the results of this subsection were established only for linear prices, and may not hold in the presence of two-part tariffs and general non-linear pricing schemes. For an excellent survey of non-linear pricing see Wilson (1993).

$$p(n, n) + (n/k)(dp/dn) - c = 0.$$

Clearly the marginal revenue for firm j lies between monopoly and perfect competition, as seen in Figure 4. Thus, *the oligopoly equilibrium network size lies between the perfectly competitive size and the size chosen by a monopolist who influences expectations*. From a social welfare point of view, the equilibrium of oligopolists that influence expectations is more inefficient than the perfectly competitive outcome but more efficient than the choice of a monopolist who influences expectations.

3. <u>Dynamics and Durable Goods</u>

To describe more accurately the FAX market, we develop next a model of durable goods competition with network externalities. The problem is more complex, since now firms and consumers need to predict accurately the whole path of future network sizes and prices. Nevertheless, we are able to show a one-to-one correspondence between a dynamic durable goods problem under perfect competition to a single-period problem.

Let the instantaneous utility of owning the network good for consumer of income y and network of size n be given by $u(y, n)$. Assume for simplicity that once a good is purchased, it yields an infinite stream of future utility. Given an expected future time path of network size $n^e(t)$, the present value of a machine purchase at time t for a consumer of income y is given by

$$V(y, t, n^e(t)) = \int_t^\infty e^{-\rho s}\, u(y, n^e(s))ds,$$

where ρ is the discount rate. If the durable good is purchased at time t at price $p(t)$, the present value of its cost is

$$q(t) = e^{-\rho t}\, p(t).$$

The consumer of income y buys at time t^* that maximizes $V(y, t, n^e(t)) - q(t)$, i.e., he solves

$$V'(y, t^*, n^e(t^*)) - q'(t^*) = 0.$$

This expression simplifies to[9]

$$u(y, n^e(t^*)) = \rho p(t^*) - p'(t^*) \equiv \lambda(t).$$

The shadow price $\lambda(t)$ plays exactly the same role in the durable goods problem as price p plays in the single period problem. In the durable goods case, $\lambda(t)$ represents the opportunity cost of buying the good at t rather than $t + dt$. The first term $\rho p(t)$ measures the cost of waiting one period, assuming that the price remains the same. The second term reduces the cost of buying today by any price increase in the time increment dt. Thus, $\lambda(t)$ represents the opportunity cost of buying today rather than tomorrow. Using this re-interpretation, we can apply and extend results from the non-durable analysis to the durable good case.[10]

In the dynamic setting, a *fulfilled expectations equilibrium* (*rational expectations equilibrium*) is a pair of paths of prices and sales $\{p(t), n(t)\}$ such that expectations are fulfilled and supply equals demand at every period, i.e., it fulfills:

demand:	$n_D(t) = 1 - G(p^e(t)/h(n^e(t)),$
supply:	$p(t) = c(t, n_S'(t)),$
fulfilled expectations of sales:	$n(t) = n^e(t),$
fulfilled expectations of prices:	$p(t) = p^e(t),$
market clearing:	$n_D(t) = n_S(t) = n(t),$

[9] $V' - q' = -e^{-\rho t}u(\omega, n^e(t)) + \rho e^{-\rho t}p(t) - e^{-\rho t}p'(t) = 0$, so that $u(\omega, n^e(t)) = \rho p(t^*) - p'(t^*)$.

[10] In particular, given instantaneous utility $u(y, n^e) = yh(n^e)$, the marginal consumer at time t is $y^* = \lambda(t)/h(n^e)$, and therefore the demand at time t is $n(t) = 1 - G(\lambda(t)/h(n^e(t))$. At a fulfilled expectations equilibrium, $n^e(t) = n(t)$, so that $n(t) = 1 - G(\lambda(t)/h(n(t))$ or equivalently $\lambda(t) = h(n(t)G^{-1}(1 - n(t))$.

where $c(t, n_S'(t))$ is the marginal cost at time t which may depend on the size of output $n_S'(t)$ at t. In the next section, we apply this dynamic analysis to the fax market in the U.S.

4. <u>The Growth of the U.S. Market for Facsimile Machines</u>

As already described in the introduction, the market for facsimile machines in the U.S. exploded during the mid-to-late 80s, with growth rates of the number of units shipped exceeding 150% in 1987. We argue that this tremendous surge in demand was not driven as much by outside shifts in consumer demand and price reductions as much as it was driven by the "feedback" effect induced by both past increases and anticipated future increases in the size of the installed base. The anecdotal evidence is consistent with this interpretation since the most dramatic fall in prices occurred well before 1987. What *is* true about 1987, however, is that this is the year in which the *rate* at which prices were falling began to taper off. This is an important clue because in the consumer's solution to the dynamic, durable goods problem, the desire to postpone a purchase is proportional to $\lambda_t = \rho p(t) - p'(t)$. This implies that as long as prices are still falling (that is, as long as $p'(t) < 0$), aggregate demand is weak. This is exactly what the data seem to show.

We now formalize the above intuition with a simple calibration exercise. In addition to our data on the average prices and quantities of facsimile machines sold in the U.S. between 1979 and 1992, empirical estimation of our model requires data on the distribution of consumer characteristics. Ideally, these characteristics would be identified by collecting marketing data on consumers that purchase fax machines and then using these data to estimate a discrete choice model. In practice, however, access to such data is difficult, so we pursue an alternative strategy that is feasible with available data. First, even though most fax purchases are made by firms and not consumers, and that many firms purchase more than one machine. We argue that it is nonetheless reasonable and convenient to model the unit demand for fax machines as a function of consumer characteristics. This is because a firm's demand for fax machines ultimately is derived from "employee demand." For example, a firm with a high fraction of highly skilled white collar workers will have a higher demand for fax machines than a firm with a high fraction of production line workers. For simplicity and feasibility, we assume that the employee characteristics related to fax demand can be summarized by employees' income.

In order to characterize the distribution of consumer types as a function of consumer income, we use data on the distribution of income from the Current Population Survey (CPS) for survey years 1976, 1981, 1986, and 1991. Since income is approximately lognormal, we transform the data using natural log to obtain normally distributed log-income. We then calculate the mean and variance for each of the above four years and then interpolate to estimate the distribution of log-income for each year between 1979 and 1992.[11] This gives us a time-varying estimate of the distribution function described in section 3, that is, $G(y; t)$. We use the notation μ_t and σ_t to denote the mean and standard deviation of log income, and the notation $\Phi(x)$ to denote the standard normal distribution. This notation allows us to represent our empirical estimate of G by

$$G(y; t) = \Phi((\ln(y) - \mu_t)/\sigma_t).$$

Next we normalize the size of the fax network by assuming that maximum potential network size is 20 million fax machines. This number implies a maximum ratio of about one fax machine for every five workers in the U.S. (we estimate the number of fax machines in 1992 to be about 7.6 million). In the results reported below, we experimented with both larger and

[11] To be specific, this procedure also requires that we account for top coding in the CPS. For each of the four years, we truncate above at \$100,000, except for 1976, which is truncated above at \$80,000. We also truncate below at \$2000 in order to avoid data problems with outliers. We then use formulas for the truncated mean and variance of a normal distribution to calculate means of log income (nominal) of 9.33, 9.38, 9.70, and 9.73 for the four years, respectively. The standard deviations of income in these four years are 1.18, 1.27, 1.31, and 1.33, respectively. The interpolation is done using a cubic polynomial. Finally we converted these numbers to real 1987 dollars using the GDP deflator reported in Table 1.1 of the Current Survey of Business. We are extremely grateful to our colleague Rick Flyer for providing us with the estimates from the CPS.

smaller values of the maximum network size and this did not affect the calibration results reported below. To construct the "stock" of fax machines, that is, the installed base, we assume that fax machines depreciate at a rate of 13.3% per year, and used a perpetual inventory method to accumulate unit sales. Here, too, we experimented with various depreciation rates and this did not significantly affect the calibration results reported below.[12] Finally, we deflated our price series for fax machines using the GDP deflator reported in Table 1.1 of the Current Survey of Business.

With the above empirical estimate of the distribution of consumer income G(y; t) and our data on normalized network size, n_t, real prices, p_t, we calibrate the model by choosing values of the remaining unknown parameters to fit the model. We simplify the model somewhat by assuming that fax machines are pure network goods, so that they give no utility in a network of size zero. That is, we assume $k = 0$. This is not strictly true, since fax machines can also double as telephones, but given the widespread availability of telephones in most places where fax machines are used, it seems reasonable to assume that the fax machines are valued only for their ability to make fax transmissions.

We relax the simplifying assumption made in the previous section that $\gamma=1$. Thus, our general Cobb-Douglas utility specification is

$$u(y_t, n^e_t) = Ay_t^{\gamma}n_{t-1}^{\alpha}.$$

[12] We chose the depreciation rate of 13.3% by applying the average service life of telephones (7.5 years), which is estimated using life expectancy tables for consumer possessions used by insurance adjusters in responding to claims for fire and theft damage. We are grateful to Peter Klenow for providing us with this estimate.

Note that we have made the assumption $n^e_t = n_{t-1}$, that is, our empirical specification assumes that the expected size of the network this year is a linear function of the network size at the beginning of the year.[13]

Recalling our use of the notation $\lambda_t = \rho p_t - p_t'$, we construct a data series for λ_t by assuming $\rho = 0.2$. Our results in the previous section show that the value of the marginal consumer is calculated by setting utility equal to λ_t and solving for y_t. Taking natural logs of the resulting expression yields

$$\ln y_t = \gamma^{1}(\ln \lambda_t - \alpha \ln n_{t-1} - \ln A)$$

Using our empirical estimate of the distribution of consumer income, the equilibrium network size is given by

$$n_t = 1 - \Phi((\ln y_t - \mu_t)/\sigma_t).$$

Inverting Φ and solving this expression for $\ln(y_t)$ yields

$$\sigma_t\Phi^{-1}(1 - n_t) + \mu_t = \ln y_t.$$

Since the inverse of the cumulative standard normal is easily calculated, and since σ_t, n_t, and μ_t are all variables in our data, the term on right side of the above expression is a variable that we construct. We define this variable using the notation $g_t = \sigma_t\Phi^{-1}(1 - n_t) + \mu_t$. Finally, substituting our expression for the marginal consumer yields our estimating equation

$$g_t = \beta_0 + \beta_1 \ln\lambda_t + \beta_2 \ln n_{t-1} + e_t$$

[13] Imposing a coefficient of one is arbitrary and reflects the fact that the constant term A absorbs this scaling factor in any case.

where $\beta_0 = -\gamma^{1}\ln A$, $\beta_1 = \gamma^{1}$, $\beta_2 = -\gamma^{1}\alpha$, and e_t is an error term that represents approximation errors in the functional form assumptions as well as errors in the measurement of g_t.

We estimate this specification of the model using OLS. We point out that this is essentially a demand equation in which a nonlinear transformation of the quantity variable appears on the left side of the equation and a price term (λ_t) and a demand shifter (n_{t-1}) appear on the right side of the equation. This interpretation exposes a potential econometric problem with the use of OLS to estimate the model. The problem is that the error term e_t could also contain unexplained variation in g_t due to surprises in the realization of prices (i.e., realizations of λ_t). That is, the price term is endogenous. For this reason, we also report results estimated with generalized method of moments (GMM) using lag values of $\ln(\lambda_t)$ and $\ln(n_t)$ as instruments.

We also experimented with various *ad hoc* modifications of the specification to assess the robustness of our calibration estimates. Table 1 reports the estimates of the model for several variations of the basic specification described above. For Model 1, the estimates reveal a large positive coefficient on the price term, as predicted, and a large negative coefficient on the network term, also as predicted by the model. Both coefficients (as well as the coefficient estimates reported for models 2 through 6) are estimated with tight standard errors, although we hasten to emphasize that these standard error estimates are grossly misleading given the very small size of our sample. We include them merely to indicate that the parameter estimates do a very satisfactory job of matching the data. We note that goodness of the calibration fit is also revealed by the high R^2 value of 0.902.

The coefficient estimates for models 2 through 6 largely confirm the results for model 1, and reveal them to be fairly robust to alternative assumptions. The estimates for model 2 reveal that the estimates (particularly the coefficient on the price term) are robust to the use of instrumental variables. In model 3 we included the lagged value of the price term (rather than

Table 1: Calibration Estimates.

Variable	Model 1 OLS	Model 2 GMM	Model 3 OLS	Model 4 GMM	Model 5 OLS	Model 6 GMM
Constant	-0.501 (0.050)	-0.527 (0.031)	-0.632 (0.051)	-0.633 (0.034)	-1.401 (1.386)	-1.103 (0.964)
$\ln \lambda_t$	0.118 (0.023)	0.100 (0.017)	...	...	...	...
$\ln \lambda_{t-1}$	...	...	0.142 (0.020)	0.153 (0.015)	0.129 (0.032)	0.144 (0.023)
$\ln n_{t-1}$	-0.582 (0.034)	-0.619 (0.022)	-0.574 (0.026)	-0.555 (0.016)	-0.616 (0.081)	-0.581 (0.055)
Year	...	...	...	...	0.008 (0.015)	0.005 (0.011)
# Obs	15	14	14	14	14	14
Adj. R^2	0.902	0.914	0.951	0.951	0.947	0.944

Note: Asymptotic standard errors appear in parentheses and are consistent in the presence of conditional heteroskedasticity. The instruments used for the GMM estimates in models 2, 4, and 6 are two lags each of $\ln \lambda_t$ and $\ln n_t$ plus the mean of log income and a time trend. Hansen's test of the over-identifying restrictions generated p-values of 0.363, 0.540, and 0.324, respectively.

the current) as an alternative means of controlling for the endogeneity of price. In particular, since one component of this term is the expected price change, one could argue that the lagged value is a good proxy for the expected value. As we expected, this slightly increased the coefficient on the price term, but the magnitude of the increase is not dramatic, nor did it change by much when we estimated the model using instruments (model 4).

Finally, model 5 includes a time trend to control for the possibility of omitted variables such as trends in the distribution of consumer characteristics that are not well-proxied by income, or changes in the quality of fax machines over time. The inclusion of a time trend somewhat reduces the coefficient on price, but again the magnitude is small, the standard error of the trend coefficient is relatively large, and the magnitude of the adjusted R^2 actually falls slightly. Hence, we conclude that there is little evidence that the price term and network term are contaminated by omitted variable bias, or that they do not do a good job of matching the model to the data.

The virtue of our "structural" approach to the specification of the empirical model is that we can interpret the coefficients on the price and network variables in terms of the preference and technology parameters of the model. In particular, the estimates in Table 1 allow us to identify values of the model parameters α and γ. For example, the model implies that the coefficient on $\ln(\lambda_t)$ is the inverse of γ. Hence, our estimates imply values of γ that range from 6.5 to 10. These estimates are not of any particular interest, except that they allow us to infer a value for α. In particular, the model implies that the coefficient on $\ln(n_t)$ is minus the inverse of γ times α. Hence, our estimates imply values of α that range from 3.6 to 6.2.[14]

5. <u>Conclusion</u>

In this paper, we discussed the equilibrium size of networks under alternative market structures for both non-durable and for durable goods. In the presence of network externalities, we showed that, for high marginal costs, the size of network is zero; as costs fall, the network size abruptly increases to a positive and significant size (the critical mass) and thereafter it increases gradually as costs continue to fall.

We generalized these results to a dynamic multi-period setting and to durable goods. In this framework, the abrupt increase of the network from zero to critical mass of the single-period

[14] These estimates imply very large network effects in our model, and they are somewhat troubling if taken at face value because they actually imply that consumer utility is *convex* in the size of the of the installed base. But upon further reflection and investigation this turns out not to be such a problem. When we simultaneously include both current and lagged values of the price term (not reported in the table), the sum of the price coefficients is 0.21, and the coefficient on the network effect drops to 0.50. This still implies an exponent of $\alpha = 2.5$ in the utility function for the installed base, but it suggests that future extensions and generalizations of the model might lead to more reasonable estimates. This is the subject of our current and ongoing research.

model is replaced by a continuous but steep increase in network size. We applied our model to US FAX market. Calibration of our model for this market suggests that its growth was strongly influenced by network externalities.

References

Antonelli, Christiano, (1992), *The Economics of Information Networks*, North Holland: Amsterdam.

Arthur, W. Brian, (1988), "Self-Reinforcing Mechanisms in Economics," in Philip W. Anderson, Kenneth J. Arrow and David Pines (eds.) *The Economy as an Evolving Complex System*, Addison Wesley.

Arthur, W. Brian, (1990), "Positive Feedbacks in the Economy," *Scientific American*, February 1990, pp. 92-99.

Cabral, Luis, (1990), "On the Adoption of Innovations with 'Network' Externalities," *Mathematical Social Sciences*, vol. 19, pp. 229-308.

Cabral, Luis, and Antonio Leita, (1989), "Network Consumption Externalities: The Case of Portuguese Telex Service," in *The Economics of Information Networks*, Cristiano Antonelli (ed.), North Holland, 1992, pp. 129-40.

Economides, Nicholas, (1994), "The Economics of Networks," plenary address, EARIE conference, September 1994, Discussion Paper no. EC-94-24, Stern School of Business, N.Y.U.

Economides, Nicholas, (1995), "Network Externalities, Complementarities, and Invitations to Enter," forthcoming *European Journal of Political Economy*.

Economides, Nicholas and Charles Himmelberg, (1994), "Critical Mass and Network Size," mimeo.

Farrell, Joseph and Saloner, Garth (1985), "Standardization, Compatibility, and Innovation," *Rand Journal of Economics*, vol. 16, pp. 70-83.

Hayashi, Koichiro, (1992), "From Network Externalities to Interconnection," in *The Economics of Information Networks*, Cristiano Antonelli (ed.), North Holland: Amsterdam.

Katz, Michael and Carl Shapiro, (1985), "Network Externalities, Competition and Compatibility," *American Economic Review*, vol. 75 (3), pp. 424-440.

Rohlfs, Jeffrey, (1974), "A Theory of Interdependent Demand for a Communications Service," *Bell Journal of Economics*, vol. 5, no. 1, pp. 16-37.

Wilson, Robert, (1993), *Nonlinear Pricing*, Oxford University Press.