PRIVATE NETWORKS PUBLIC OBJECTIVES
E. Noam and A. NíShúilleabháin (Editors)

Network Evolution and Coalition Formation[1]

Graciela Chichilnisky

1. INTRODUCTION

International financial networks such as Euroclear, CEDEL, SWIFT, and Global Custody networks, are important factors in the rapid development of cross-border securities trading.[2] The value of the transactions going through these networks is high: more than US $4 billion worth of cross-border equities are transacted daily.

Network services exhibit two distinctive characteristics: (1) The willingness to pay increases with the level of activity, because the network is more valuable when it has more uses;[3] and (2) Unit costs decrease with the volume sold, because there are large fixed set-up costs.[4] Because of these characteristics the economic fundamentals of the market for network services can be opposite to what is predicted in classic textbook examples: the supply can be downward sloping, while the demand can increase with prices.[5] For these reasons the evolution of the market for network services can be very different from that of standard markets.

As global capital markets evolve the communications and administration of trading requires more complex and extensive networks. The emergence of the services needed to support global capital markets depends on the feasibility or the survival of the network service. A large "capital mass" of users may be necessary before a network producer "breaks even" i.e. achieves positive profits.[6] Once this critical mass is achieved, however, there are increasing profits to be obtained from each additional user. For this reason setting-up a network is often described as a "start-up" problem. This problem can be difficult to solve: Many potentially valuable networks fail to emerge or to survive. This paper argues that the formation of certain coalitions of users can solve this problem. The idea is that if no one can find smaller coalitions of users which produce strong positive externalities to each other, the size of the critical mass can be reduced, and the start-up problem can be overcome. The literature has not examined so far the formation of coalitions of users which can exploit the externalities which they induce on others. This is the main focus of the paper.

1.1. Network Coalitions

The formation of coalition of users can be crucial for the network's operation. Many international securities networks are organized into "clusters" of users (coalitions) who "communicate" with each other much more frequently than they do with the rest of the users. Examples are global custodians cross-border securities holdings. They use a network such as SWIFT and CEDEL to make and receive payments or other instructions with several thousand entitites across the world. In addition, these banks communicate routinely with about forty other banks worldwide, which are called their "subcustodians." Subcustodians handle the paper instruments, and administer and report on taxes and corporate actions in each country. Furthermore, the global custodians also have a network of clients, several hundreds on occasions, with whom they also communicate very frequently. Each communication is about a cross-order transaction of typically US $50,000 or more.

The bank has, therefore, two levels of network use: "infrequent use" to communicate with a large number of institutions across the world, and "frequent use" to communicate with a smaller group of institutions such as subcustodians or customers. The ability to communicate with the latter, i.e. at the "frequent use" level, is very valuable. Fewer parties of the second type are needed in the network for the user to reach the same level of benefit that the user derives from infrequent use. Therefore, if a cluster of users of the second type become network users, in practice the "critical mass" required to break even can be smaller.

1.2. The Network Evolution

With these practical applications in mind we formalize a network market with many users. Using game theory and dynamic stochastic analysis, we show how the network evolves. We define a critical mass, define stochastic process of coalition formation through time, and specify the long-run properties of the resulting network market.

We explore the formation of coalitions of users when the externalities produced by the players are heterogeneous, i.e. when there exist clusters of players which produce more externalities to each other than they do to the rest of the potential users. Proposition 12 and Corollary 13 establish that the gains from decentralization in this context, i.e. the gains from distinguishing those clusters and producing a cluster of networks rather than one big centralized network, are surprisingly large. We show below that the probability of success of the network *increases exponentially* with decreases in the size of the clusters. This may account for the actual network structure (clusters of users) that one observes in practice (e.g. global custody networks).

Proposition 12 and several examples explore the gains to decentralization formulated by calculating the stopping time until coalitions of critical mass are formed. This is financially important since until critical mass is reached, profits are negative. The critical mass of the network (Section 5) measures its economic feasibility in terms of the number of players which are required for positive profits and determines the number and the stability properties of Nash equilibrium of a network game (Sections 3 and 4, Propositions 8 and 9 and Corollaries 10 and 11). We show that the set of Nash equilibria are quite different under different information structures.

1.3. A Dynamic Network Game

With these applications in mind, we formalize the network market and study a dynamic game determining its dynamics and equilibria or steady states.

Users come into the network following a stochastic process. They may stay or leave depending on the economic incentives. There are large and small users; the former are informed about the externalities which they produce to other users, and the latter are not. Two scenarios are considered. In the first, the externalities between the users are homogeneous: all players within a certain group must simultaneously join the network in order for the critical mass to be achieved. A second scenario studies the expected length of time required to reach a critical mass of users with clusters of users which are heterogeneous in terms of the externalities they produce to each other. We prove the existence of solutions and the number of solutions under different characteristics of the users. We explore the characteristics of the critical mass and the difficulties of the start-up problem. We show in Proposition 12 that the probability of success (survival) of the network increases exponentially with decreases in the size of the clusters. Somewhat suprisingly, while the size of the clusters is all important, the number of clusters required to break even is almost irrelevant.

2. THE ECONOMICS OF INTERNATIONAL FINANCIAL NETWORKS

We focus on two main characteristics of international financial networks:

Communications externalities: The parties exchange information through the network. Thus the more parties that are accessible through the network, the more valuable the network's communicating ability is.

Audit trails: A historical record of each transaction must be kept. Audit trails are records of the messages sent, by whom, and when, and of the actions taken by the different parties with respect to each trade.

These characteristics motivate the following definitions:

Let $u_i(q_i,x)$ be the i th user's utility from consuming a quantity q_i of the network services or messages, and let x be the number of users, indexed by the integers. The variable q_i is either 0 or 1 depending on whether the network is used or not. Since the utility derived from using the network increases with the number of other users:

$$\frac{d}{dx}u_i(q_i,x) > 0 \tag{1}$$

Similarly, when there are no users the utility of using the network is zero:

$$u_i(q_i,0) = 0 \tag{2}$$

Assume all users use the same amount of network services, and choose units of measurement so that the total quantity of network services consumed is equal to the number of users.[7] Let $u_i = u_j$ for all $i, j \in X = R^+$. Conditions (2.1) and (2.2) imply that users are willing to pay higher fees for the same network services when the network has more users, and that at zero network use, they will only wish to pay 1. Formally:

$$\frac{d}{dx}D(x) > 0 \tag{3}$$

and

$$D(0) = 0$$

where $D(x)$ is the x's user "willingness to pay" for the network services, given that up to x users are already using the network. The associated demand function is also denoted D.

In economic terms the network's data bases and switches are *fixed costs* in the provision of network services, since they must be incurred independently of the amount of network use. These costs are generally quite large and typically incurred once. This implies that the average cost of a message decreases with the number of messages, so that there are increasing returns to scale in the production of the network services. The average cost of a message typically decreases and goes to zero as the number of messages goes to infinity, which we now assume. This gives rise to an average cost curve denoted $C(x)$, satisfying:

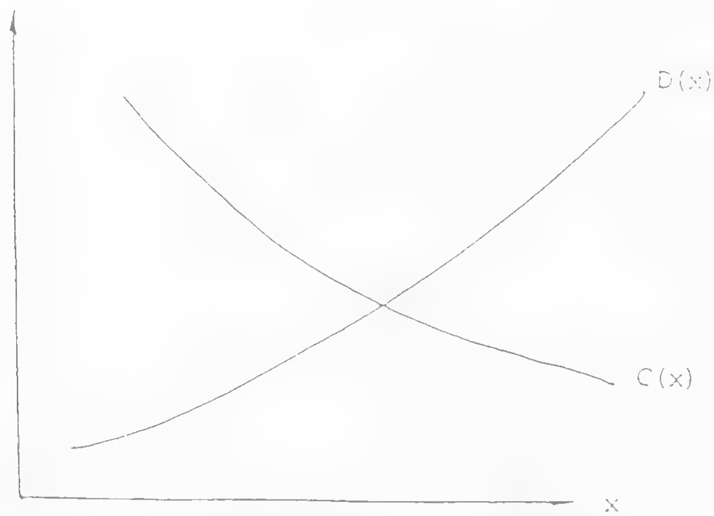
$$\frac{d}{dx} C(x) < 0 \qquad (4)$$

and

$$\lim_{x \to \infty} C(x) = 0 \qquad (5)$$

We assume $C(x)$ that is continuous. The associated supply curve is denoted $S(p)$, where p denotes price. The user's externalities lead to an upward sloping demand curve $D(x)$, while increasing returns in production lead to a downward sloping average cost curve $C(x)$. Diagram 1 illustrates.

Diagram 1



D(.) is willingness to pay; C(.) is average cost.

We are concerned here with the "start up problem"; therefore we assume that a producer's main concern is to break even in order to cover its fixed costs and operate with non-negative profits. An *Average Cost Pricing Equilibrium* is defined as a price-quantity vector at which the market clears, and producers charge at average cost. At such an equilibrium, producers break even:

***Observation**:* Under conditions 2.1, 2.2, 2.3, 2.4, and 2.5, there is a unique market clearing average cost pricing equilibrium (p^*, x^*) such that x satisfies $D(x^*) = C(x^*)$ and $p^* = D(x^*)$. This equilibrium is unstable under either the *quantity adjustment process*

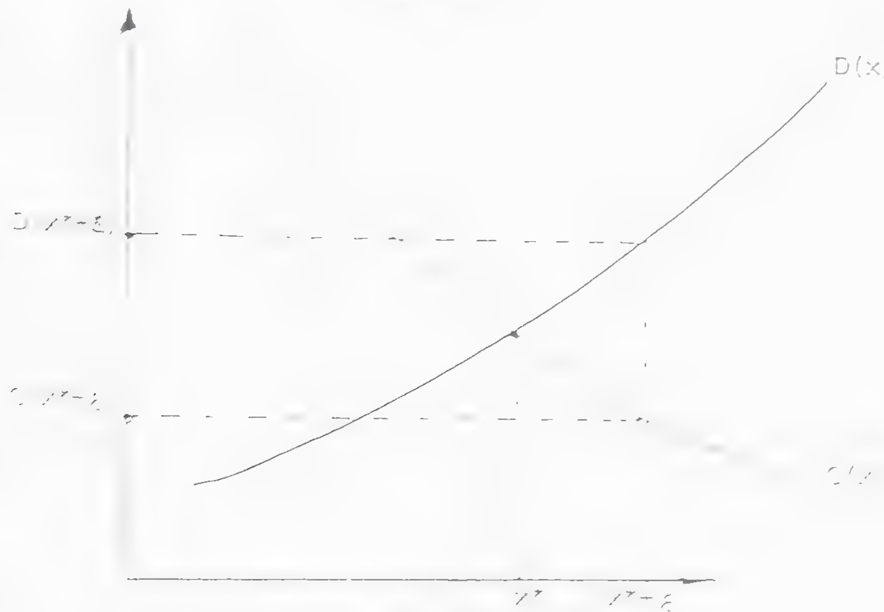
$$(2.6) \quad \dot{x} = \lambda(D(x) - C(x)), \text{ for a real number, } \lambda > 0, \qquad (6)$$

or under the *Walrasian price adjustment process*

$$(2.7)\ \dot{p} = \mu(ED(p)), \text{ for a real number, } \mu > 0, \qquad (7)$$

where $ED(p)$ denotes the excess demand function $D^{-1}(p)-C^{-1}(p)$. This market clearing equilibrium is Pareto inefficient as it undersupplies network services: for all $\epsilon>0$, (p^*,x^*) is Pareto inferior to a non-market clearing allocation (a price - quantity vector) with quantity $x^*+\epsilon$ users and with prices defined by their willingness to pay $D(x^*+\epsilon)$. Producers charge according to the average cost curve $C(x^*+\epsilon)$.

Diagram 2



x* is not Pareto efficient.

A proof is given in the Appendix.

Definition 1: A *Critical Mass* of users is the *quantity* of users at which producers break even: the willingness to pay curve intersects the average cost curve.

A critical mass of users consumes a quantity of network services which is equal to that produced at the average cost pricing equilibrium. At this quantity of network services, each user's willingness to pay equals or exceeds average cost, leading to no loss or to a net gain to the user who joins the network. Below the critical mass the opposite is true, namely average costs exceed willingness to pay, so there is a net loss from using the network.

3. THE NETWORK GAME

We shall now consider strategic moves on the part of the users. Through time, users choose strategically whether to join the network or not, or whether to leave the network once joined. Over time, a typical user will join or leave the network several times. The user's strategic decisions lead to a level of network use, and therefore to an average cost and to a price they are willing to pay.

In order to simplify notation users are now indexed by the positive integers, i.e. $X = \{1,2 \ldots\}$.

A *Network Game G* is defined as follows. The *players* are all the potential network users in the set *X*. Each player $y \epsilon X$ has *two possible strategies*: to use the network or not to use it: if $\varphi(y)$ is player y's move, then $\varphi(y)$ is either 0 or 1. Through time a player may either join or leave the network, and may do so several times. The *quantity of network use* is the sum of the player's strategies:

$$x = \sum_{y} \varphi(y) \tag{8}$$

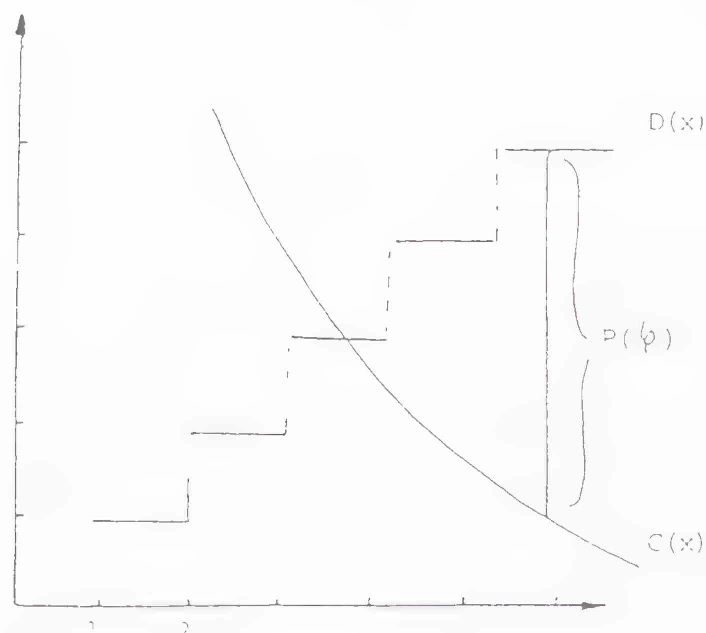
Assuming that producers charge at average costs, the *payoff* to the players who use the network are their welfare gains (or losses), defined as the difference between two prices: the average cost and the willingness to pay. In other words: a player is better off the larger its willingness to pay in relation to the average cost which it pays for the service:

Definition 2: The *payoff* to the player who plays strategy 1 is the difference between the user's willingness to pay and the average costs computed at the sum of the players' strategies:

$$x = \sum_{y} \varphi(y),\ i.e. \tag{9}$$

$$P(\varphi) = D\left(\sum_{y} \varphi(y)\right) - C\left(\sum_{y} \varphi(y)\right) \tag{10}$$

Diagram 3



The payoff to a player joining the network when the game allocation is φ.

The payoff to the player who plays the strategy zero is zero, i.e. $P(0)=0$.

All players know the average costs and demand function, as well as the total number of users.

Definition 3: We define a large player as one who is aware of the impact of its strategy on prices, and acts accordingly.

Definition 4: A small player is one who acts as if it had no influence on prices.

Each player aims at maximizing payoffs. Denote the cardinality of the set of players X by X_{max}, with $x^* < X_{max} < \infty$.

Definition 5: A game allocation[9] φ^* is a Nash equilibrium when, for each player x', $\varphi^*(x')$ is the optimal response of player x' to the strategies played in φ^* by all the others.[10]

Definition 6: An adjustment process is defined now as follows:

$$\varphi_{t+1} = \varphi_t^+ \tag{11}$$

where φ_t^+ is the optimal reaction to allocation φ_t, defined so that for each player x:

$$P(\varphi_t^+(x), \varphi_t^0(x)) = \max_{\theta} P(\theta, \varphi_t^0(x)) \tag{12}$$

Observe that a Nash Equilibrium is a *steady state* of this adjustment process, i.e. a game allocation φ^* such that $(\varphi^*)^+ = \varphi^*$.

Definition 7: The *region of stability of a steady state* φ^* denoted $S(\varphi^*)$, is the set of all game allocations φ^0 such that if the adjustmant process start at one such allocation, then it will converge to the steady state φ^*. Formally $\varphi^0 \in S(\varphi^*)$ iff $\varphi_{t=0} = \varphi^0$ implies:

$$\lim_{t \to \infty} \varphi_t = \varphi^* \tag{13}$$

4. NASH EQUILIBRIA WITH SMALL AND LARGE PLAYERS

The steady states of the network are quite different when the traders are small as compared to when they are large. The following results are proven in the Appendix.

Proposition 8: When players are large, the network game G has two Nash Equilibria, denoted E_0 and E_{Xmax}: the former has no network use, and the latter includes all users. Both equilibria are locally unique, and they are also locally stable according to the adjustment process (3.1). Once the usage of the network up to the critical mass minus one is reached, the adjustmant process converges to the equilibrium E_0, so that the network does not survive. Once usage on or above the critical mass is achieved, this leads necessarily[11] to the equilibrium E_{Xmax}. The equilibrium E_0 is not Pareto efficient. The equilibrium E_{Xmax} is Pareto efficient, but at E_{Xmax} the market does not clear; there is an excess demand for network services.

Proposition 9: Assume that players are small, and that at the critical mass usage the costs equal the willingness to pay. Then the network game G has as many Nash equilibria as the number of all combinations of players into subgroups of critical mass size, plus two i.e. $(Xmax_{x^*}) + 2$. The latter two are the Nash equilibria E_0 and E_{Xmax} defined in Proposition 8. The Equilibrium E_0 is Pareto inferior. Once the number of users reaches the level of critical mass minus one, the network will converge to the equilibrium E_0. E_{Xmax} is a Pareto efficient and stable Nash equilibrium, but the market does not clear at E_{Xmax}. Once the number of users equals or exceeds the critical mass plus one the network converges to the equilibrium E_{Xmax}. All the $(Xmax_{x^*})$ equilibria with a critical mass are unstable. Furthermore, if at the critical mass the willingness to pay exceeds the network costs, then there are only two Nash equilibria, E_0 and E_{Xmax}.

5. CRITICAL MASS

The results presented so far indicate the importance of the critical mass in determining the network's behaviour. The critical mass is the smallest number of users at which producers' profits cease to be negative. Below the critical mass, the network is not sustainable since firms make negative profits. The following corollaries show a negative correlation between the size of the critical mass and the ability of the network to converge to its Pareto Optimal position. They also show how fixed costs and the level of externalities can influence the size of the critical mass itself:

Corollary 10: The area of convergence to the Pareto Optimal Nash Equilibrium decreases as the critical mass increases.

Proof: This follows from Proposition 9.

Corollary 11: The critical mass of the network decreases when fixed costs decrease, or when the externalities among the users increase.

Proof: This follows from conditions 2.1, 2.2, 2.3, 2.4, and 2.5 of the function *C* and *D*.

However important is the critical mass for both producers and users, it is an unstable position. If the critical mass is exceeded, usage will immediately increase towards the e Pareto Optimum, E_{Xmax}. If, however, the critical mass is not reached, usage will inevitably dwindle to zero, i.e. to the Pareto inferior equilibrium E_0. These are the conclusions we reach when examining the network usage problem as a non-cooperative game for the users. In order to reach a Pareto efficient solution, it is essential to reach the critical mass. This requires the simultaneous decision by at least x^* users to join the network. A cooperative solution, namely the formation of a coalition, could resolve the problem and lead the economy towards the Pareto Optimal solution E_{Xmax}. The critical mass required for optimality, therefore calls for the formation of coalitions to resolve the network start up problem.

6. START-UP: A GAME OF COALITION FORMATION

The difficulty in reaching a Pareto efficient outcome resides in the formation of coalitions of the right size, at least a critical mass size. This is a necessary condition for the network to break even and thus to its commercial feasibility. This is called the network "start up" problem. The amount of time needed to reach a critical mass is crucial in the financial feasibility of the "start up." A common strategy[12] is to subsidize the first users until critical mass is reached and then charge according to monopoly pricing or any other feasible pricing rule. If such a strategy is followed, and *F* represents fixed costs at each period *t*, then in financial terms the network must justify a maximum loss of \$*F* per period during each time $t \geq 0$ period until a *T* is reached at which $x(T) \geq x^*$. In other words, the network may have to justify a loss of up to:

$$ML = \sum_{t=0}^{T} F \lambda^{-t} \tag{14}$$

where *T* satisfies $x(T) \geq x^*$, and $0 < \lambda < 1$ is a discount factor. Obviously, any strategy which minimizes *T* makes the start up problem easier. In particular, considering the present discounted value of a stream of net revenue over an infinite time horizon, denoted *R*, the decision problem of the network manager is whether *R* is smaller or exceeds *ML* at an appropriate discount rate λ. This problem is obviously very sensitive to the value of *T*.

The following shows how T may decrease dramatically when instead of aiming at the formation of one coalition of critical mass, we form several "locking" subcoalitions of smaller size. A possible rationale for seeking the formation of such subcoalitions is that the players are heterogeneous: they naturally divide into subgroups within which players produce stronger externalities to each other than they do to the rest of the network users. An example is provided by dividing the population of users in subsets of users which communicate more frequently with each other than with the rest of the network, or those subsets of users who share a common data base or node. Rohlfs[13] calls this phenomenon a "non-uniform calling pattern."

7. EXAMPLE 1: AN HETEROGENEOUS NETWORK WITH 6 PLAYERS

Consider a network with six users, indicated with the letters a to f. Users a, b and c form subcoalition I, and within this group users produce externalities of value 20 to each other. The same is true for users within subcoalition II, composed of users d, e, and f. The externalities produced by a member of subcoalition I to a member of subcoalition II are always equal to 2. Assume that externalities are symmetric, i.e. player i produces the same externality on player j as j does on i. The average level of user 0 to user externalities in this network is $(20\times 2+6)/5 \sim 9$. Assume for simplicity that average costs are constant $C(x)=35$. Consider now the *willingness to pay $D(x)$* as defined in Sections 1 and 2, averaged over all possible players. Formally, this is generated by the average externalities between the users so that $D(1)$=average externality to any player of one other player being in the network. Then $D(1)\sim 9, D(2)\sim 18$ etc. Then we need at least four users to form a critical mass of the network, since $D(4)=36>35$ while for any $x<4, D(x)<35$.

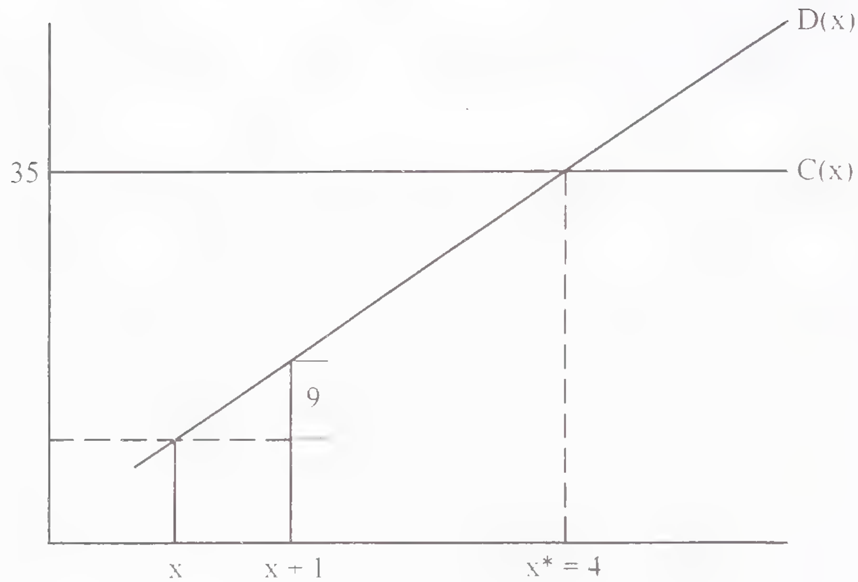
Note however that if any *two* users *within one group* agree simultaneously to use the network, their specific willingness to pay denoted $D^*(a,b)=40$ which exceeds the average willingness to pay for two players, $D(2)=18$, is larger than average costs $=35$. Thus, if two such users within one subcoalition agree simultaneously to use the network, their payoff exceeds average costs and thus in terms of the strategies defined for the game G, they will stay in the network. We indicate this by saying that these two players are "locked in." The formation of a "decentralized" critical mass, (consisting here of the two players in subcoalition I) depends then not only on the number of users but also on what users join the network at any one time. When users are heterogeneous, a critical mass could then be achieved much more quickly when there exist "clusters" of users with stronger externalities.

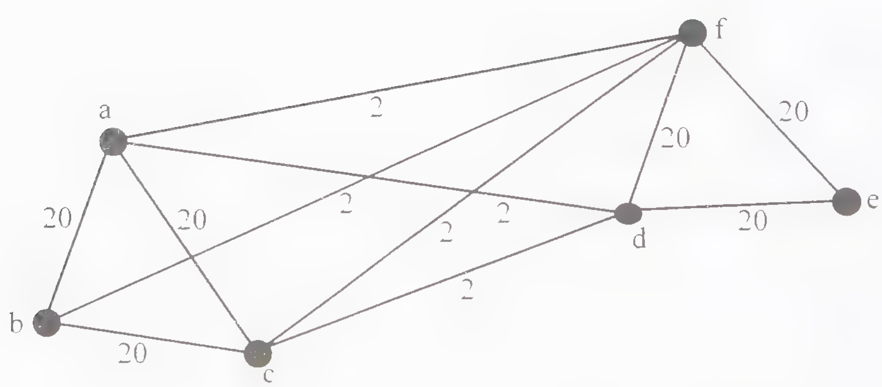
8. RANDOM COALITION FORMATION WITH HETEROGENEOUS USERS

With Example 1 in mind we seek to formalize the formation of users' coalitions by a dynamic game generated by a random process with memory. As in sections 1 and 2 there is a set X of potential player indexed by integers, of cardinality X_{max}. Since users are heterogeneous, they produce different externalities to each other. Certain subsets S_γ of β players each (called subcoalitions of critical mass) are given initially. They consist of players who produce more externalities to each other if they use the network, than they would produce to others outside their subsets. Assume that the willingness to pay of each player in the set S_γ when all others in S_γ are in the network, matches or exceeds the average costs of serving the β players in this subcoalition. Formally, let e_{ij} denote the externality produced by

Diagram 4

A Heterogeneous Network with six Players





user i on the user j which we may assume is symmetric ($e_{ij}=e_{ji}$). Consider the average externality on the j player, defined as:

$$e_j = \frac{1}{X_{max}} [\sum_{i=1,\ldots,X_{max},\, i \neq j} (e_{ij})] \quad (15)$$

This defines an average willingness to pay denoted $D(.)$, where $D(n+1)$-$D(n)=e$, or $D(n)=ne$ where e is defined by:

$$e = \frac{1}{X_{max}} \sum j\, e_j \quad (16)$$

By the assumption made on the subcoalitions:

$$S_\gamma \text{ for each j in } S_\gamma \quad \sum_{i,j \in S_\gamma} e_{ij} \geq D(x^*) \quad (17)$$

where x^* is the critical mass defined from the average costs $C(x)$, as in Section 2. There are α disjoint critical mass subcoalitions S_γ with $\alpha \times \beta = x^* \leq X_{max}$.

At any time t=0,1,2,... each of the in X players has an equal change of joining or not joining the network. The decision of each player to join the network or not at time t is independent from the decisions of all other players at time t, and is also independent from the decision of this player at other times as well *except that* if at any time t the β members of one subcoalition S_γ agree simultaneously to join the network, then their payoff meets the average costs and they are "locked in": for all $t \geq t'$ the payers in S_γ remain in the network i.e. $c(j)(t')=1 \forall t' \geq t, j \in S_\gamma$. The game continues with other players sometimes joining and sometimes leaving the network. Typically each player will join and leave the network several times. However, once locked into a subcoalition with a "critical mass," the player will remain locked in. The random game continues without the members of the subgroups who have agreed, i.e. with $k.\beta$ less players, where k is the number of subgroups who have agreed at any time t. This random game formalizes the formation of coalitions in the network.

We study the expected numbers of periods until a critical mass of players x^* have joined the network via subcoalitions S_γ (denoted ET). We compare this with the expected number of periods needed for a critical mass x^* of players to decide simultaneously (at any one time t) to join the network. The problem of finding a coalition of "average" critical mass (x^*) to make the network sustainable is then related to that of forming several smaller (decentralized) subcoalitions with total cardinality adding up to the average critical mass. We show that on average the speed with which the critical mass is reached with decentralized coalitions increases exponentially with decentralization (i.e. with the size of the subcoalitions) but it is surprisingly indifferent to the total number of such subcoalitions. This indifference is the source of decentralization efficiencies.

With this background, the random game of coalition formation is formalized mathematically as follows. At each trial α groups of β fair coins are tossed, $\alpha \times \beta = x^*$. We stop as soon as every one of the groups of coins has come up all heads in at least one trial. We call this the formation of a α,β- *decentralized coalition*, namely one made up of α subcoalitions of size β each. We compare the expected number of trials ET to form such a α,β decentralized coalition with the expected number of trials required for all coins to come up heads simultaneously. The latter is the expected number of trials required to reach a α,β-*centralized coalition*, namely one coalition with $\alpha \times \beta$ members. The comparison measures the benefits of forming decentralized vs. centralized coalitions. Obviously, the expected number of trials is lower for the $\alpha\beta$- decentralized coalition. It is less obvious, however, that the speed gained is exponentially increasing with decreases in the size of the subcoalitions, i.e. the number α, and is practically indifferent to the number β of such subcoalitions[14]:

Proposition 12: We compare the expected number of trials ET to form such a β coalitions of α traders each, i.e. an α,β- decentralized coalition, with the expected number of trials required for the formation of a coalition of size $\alpha\times\beta$, i.e. a α,β *centralized coalition*. The benefit of forming decentralized as opposed to centralized coalitions is exponentially increasing with decreases in the size of the subcoalitions, i.e. with the number α, but is practically indifferent to the number β of such subcoalitions. Formally: the expected number of periods needed to achieve a α,β decentralized coalition, ET, is approximately $2^{\beta}\log\alpha$, when α and β are large. The speed gained is exponentially increasing with decreases in the size of the subcoalitions, i.e. with the number α, but it is practically indifferent to the number β of such subcoalition (the proof is in the appendix).

9. EXAMPLE 2. GAINS FROM DECENTRALIZING A NETWORK

Consider fifteen players, divided into three disjoint sets ($\alpha=3$) denoted subcoalitions S_γ of five players each ($\beta=5$). Each player within a given subcoalition produces an externality worth eighteen to all other players in the same subcoalition. Each player in S_γ produces an externality worth one to players in other subcoalitions $S_{\gamma'}$, for $\gamma' \neq \gamma$. *On average, the willingness to pay* $D(.)$ increases by $[(18\times4)+10]/15=82/15=5.5$, with each new user joining the network. Willingness to pay for each additional user (derived from the average externality) is therefore 5.5. For simplicity assume that average network costs $C(.)$ are constant $C(x) \equiv 71.8$. Since $14\times5.5\sim77$, and $13\times5.5\sim71.5<72$, the average critical mass x^* required by the network to break even is 15 players, i.e. $x^*=X_{max}$. However, a subcoalition of all five players within one of the groups S_γ produces sufficient externalities to each other to "lock in," since $18\times4=72>71.8$. As these five players pay the average costs the network breaks even. Each of the five users has a positive payoff, and thus an incentive to stay in the network.

Consider now the expected number ET of trials until all three five player subcoalitions join the network, which is the Pareto efficient solution $E_{X_{max}}$ of Propositions 8 and 9. By Proposition 12 this number is $ET\sim2^5\log3\sim59$. In contrast, the expected number of trials for an average critical mass of users (15) to join the network simultaneously, is $2^{15}\sim32{,}768$.

The following corollary formalizes the remarkable gains from following a decentralized approach to coalition formation:

Corollary 13: The ratio of the expected number of trials to reach critical mass with a α,β decentralized coalition to that required to reach it with $\alpha\times\beta$ centralized coalition is $\log\beta/2^{\beta(\alpha-1)}$.

10. PREVIOUS LITERATURE

While networks and their critical mass have been analyzed before, the literature has not examined so far the formation of coalitions of users which form to exploit the externalities which each user produces to the others. This is the main focus of the paper. The second focus is to explore heterogeneous externalities, and how exploiting them increases dramatically the economic feasibility of the network. Related literature on networks include Rohlfs,[15] Oren and Smith,[16] Katz and Shapiro,[17] Farrell and Saloner,[18] and Heal.[19] All of these works focus on different problems and look at them from different angles than ours. Oren and Smith[20] examine the critical mass issue but they occur on the effect of different pricing structures under the assumption that users maximize benefit minus costs and a monopoly supplier maximizes profits. Katz and Shapiro[21] develop an oligopoly model in which consumers value "compatibility" between products, which they call network externalities. They study a different set of problems: the social and private incentives for firms to produce compatible products, or to switch between compatible and incompatible products. Similarly, Farrell and Saloner[22] study the problem of benefits to consumers and firms from standardizing product. Heal[23] studies Nash equilibrium usage patterns of networks and their stability properties. All these works have some points in common with ours, because they consider user's externalities, the critical mass problem, and the non-cooperative equilibria strategies of users in joining or leaving a network. However, with the exception of Heal[24] none of these pieces examine the number of equilibria, nor their welfare properties in terms of Pareto efficiency. Nor do they consider the global stability properties of such games. None of these pieces, including Heal's,[25] study how the solutions vary with the size or knowledge of the players as is done here. Finally, none of these works analyzes the formation of coalitions involved in the "startup" or feasibility problem with heterogeneous users, nor the speed of convergence to a solution, which we do in order to compare centralized and decentralized networks. The interest of our formalization and results was anticipated in Rohlfs,[26] who proposes that it would be important to explore the "startup" problem as well as the differences introduced in its solution in networks with "non-uniform calling pattern."

11. APPENDIX

11.1. Observation in Section 5

Proof: This observation is easy to establish. Conditions 2.1, 2.2, 2.3, 2.4, and 2.5 imply that the intersection of the two curves $D(x)$ and $C(x)$ exists and is unique. The slopes of $C(x)$ and $D(x)$ imply that at any quantity lower than x^*, average cost prices exceed user's willingness to pay, leading to a drop in production under the quantity adjustment process (2.6). At quantities exceeding x^* the willingness to pay exceeds average cost prices, leading to a tendency to increase output under the same adjustment process. Similar arguments are used to prove instability under the Walrasian price adjustment process in (2.7). It is immediate that (p^*, x^*) is Pareto inferior to $(D(x^*+\epsilon), x^*+\epsilon)$: this follows from the properties of the demand and the average cost supply curves, since user's utilities increase with the difference between their willingness to pay and the average cost.

11.2. Proposition 8

Proof: Consider the game allocation E_0, where each player plays the 0 strategy, i.e. $\varphi(y)$ for all y in X. For any given player y_0, $0=P(0,\varphi^0(y_0)) > P(1,\varphi^0(y_0))$ because $P(1,\varphi^0(y_0))<0$ since:

$$\sum_{y} \varphi(y) < x^* \quad (18)$$

so that average cost exceeds willingness to pay. Since this is true for all y_0 in X, E_0 is indeed a Nash Equilibrium. This Nash equilibrium is stable in the area of allocations where the sum of player's strategies does not exceed x^*-1. This is because at any game allocation φ with:

$$\sum \varphi(y) < x^* - 1 \tag{19}$$

a player who is not in the network receives a zero payoff, which cannot be improved by this player joining the network since, in the latter case, payoffs are negative or zero. On the other hand, a player who uses the network can increase its payoff by playing the 0 strategy instead, thus increasing its payoff from a negative number to zero. E_0 is therefore a Nash equilibrium and its region of stability consists of those allocation φ satisfying:

$$\sum \varphi \leq x^* - 1 \tag{20}$$

E_0 is Pareto inefficient. Consider now the network allocation E_{Xmax} in which all players play the strategy 1, i.e., $\varphi(y)=1$ for all y. E_{Xmax} is clearly a Nash equilibrium: since by assumption $x^*<X$, every player will play strategy 1 when all others use the network, as this leads to the maximum possible payoff, namely $D(X_{max})$-$C(X_{max})$. Consider now any allocation φ where the sum of all player's strategies is larger than or equal to x^*. Then at this allocation the optimal strategy for any player who is not in the network (and therefore receives a zero payoff) is to join the network. Since payoffs are positive for network usage above x^* and $x^*<X_{max}$, by joining the network this player increases its payoff to a positive number. Similarly, a player who uses the network at the allocation φ cannot improve the payoff by playing the zero strategy, because a payoffs are larger than or equal to zero when the network usage is at or above x^*. This shows that the area of stability of the equilibrium E_x contains all allocations with φ with:

$$\sum \varphi(x) \geq x^* \tag{21}$$

It is immediate to see that the Nash Equilibrium E_x is Pareto optimal, since at E_x the maximum payoff $D(X_{max})$-$C(X_{max})$ is achieved.

11.3. Proposition 9

Proof: The allocations E_0 and E_{Xmax} are both Nash equilibria: the proof is the same of that of Proposition 8. The regions of stability have now decreased, because small players do not believe that their use of the network will alter the price/willingness to pay relation, and therefore the payoffs. Any allocation where the network is used by a set of players of critical mass size, is now a Nash equilibrium. This follows from the fact that players are small: if a player uses the network when the total number of users equals the critical mass, then it does not use the network, and has therefore a zero payoff, it will not gain by joining the network because at the critical mass the payoff is the same, i.e. zero. Consider now the case where, at the critical mass, payoffs are positive.[27] Then no allocation where the numbers of payers equals the critical mass is a Nash equilibrium. This is because at such an allocation, the payoffs of joining the network are larger than those for not joining it for any player who is not already using the network.

11.4. Proposition 12

Proof: Define for any fixed $n>1$, the event A_i={The coins in group i have not come up heads in any of the first n trials}. Then:

$$T>n = \cup^{\alpha}_{i=1} A_i \tag{22}$$

$$P(A_i) = (1-(\frac{1}{2})^{\beta})^n \tag{23}$$

$$P(T>n) = \sum_i P(A_i) - \sum_{i<j} P^2(A_i) + \sum_{i<j<k} P^3(A_i)\ldots \tag{24}$$

Letting $c=1\text{-}(1/2)^{\beta}$ we have:

$$P(T>n) = \binom{\alpha}{1} c^{n} - \binom{\alpha}{2} c^{2n} + \binom{\alpha}{3} c^{3n} \tag{25}$$

$$ET = 1 + P(T>1) + P(T>2) + \ldots \tag{26}$$

$$= 1 + \binom{\alpha}{1}[c+c^2] - \binom{\alpha}{2}[c^2+c^4]\binom{\alpha}{3}[c^3+c^6] \tag{27}$$

$$= 1 + \binom{\alpha}{1}\frac{c}{(1-c)} - \binom{\alpha}{2}\frac{c^2}{(1+c^2)} + \binom{\alpha}{3}\frac{c^3}{(1-c^3)} \quad \ldots \tag{28}$$

or:

$$ET = 1 - (2^{\beta}-1)[\binom{\alpha}{1} - \binom{\alpha}{2}\frac{c}{(1+c)} - \binom{\alpha}{3}\frac{c^2}{(1+c+c^2)}] \tag{29}$$

If β is large then $c>1$, and

$$ET \sim 2^{\beta}[\binom{\alpha}{1} - \binom{\alpha}{2}/2 + \binom{\alpha}{3}/3 \ \ldots + (-1)^{\alpha-1}\binom{\alpha}{\alpha} \ \alpha] \tag{30}$$

Now, for every $\alpha>1$:

$$\sum_{i=1}^{a} (-1)^{i+1}\binom{\alpha}{i}/i = 1+\frac{1}{2}+\ldots\frac{1}{a} \tag{31}$$

so that

$$(10.2) = 2^{\beta}[1+\frac{1}{2}+\ldots\frac{1}{\alpha}] \sim 2^{\beta}\log\ \alpha \tag{32}$$

if α is also large.

REFERENCES

Farrell, J. and G. Saloner, "Standardization, compatibility, and innovation," *Rand Journal of Economics*, Vol 16, No. 1, Spring 1985.

Heal, G.M., "Price and Market Share Dynamics in a Network Industry," Working Paper, Columbia University Graduate School of Business, New York, 1990.

Howell, M. and A. Cozzini,"International Equity Flows-1989 Edition," Salomon Brothers European Equity Research London, August 1989.

Katz, M. and C. Shapiro, "Network Externalities, Competition and Compatibility," *The American Economic Review*, Vol 75, No. 3, pp. 424-40, June 1985.

Oren, S. and S. Smith, "Critical Mass and tariff structure in electronic communications markets," *The Bell Journal of Economics*, 12, pp. 467-86, Autumn 1981.

Rohles, J., "A theory of interdependent demand for communication services," *The Bell Journal of Economics and Management Science*, Vol 5. No 1, pp. 16-37, Spring 1974.

ENDNOTES

This paper was presented in seminars at the Department of Economics, Columbia University, February 1990, at the Economics Department, Stanford University, August 6, 1990, at the Economics Department of Harvard University September 1990, at the Computer Sciences Department of Yale University November 1990, at the Starr Center New York University May 1991, and at the first meeting of the Brown-Columbia-Hopkins Workshop in 1992. Herbert Robbins and Geoffrey Heal provided valuable comments and suggestions.

[2] Cross-border securities are stocks or bonds which are held in custody or listed in the stock exchange of one country, and traded in another country.

[3] This is called a "positive externality" across users.

[4] This effect is called "increasing returns."

[5] See the discussion below.

[6] For example, McCaw Communications, which was sold recently for several US $billions to AT&T, never reached the break even point.

[7] Because of this assumption and the fact that q_i is either 0 or 1, the charges to each user are what is usually denoted "connect charges."

[8] All properties of the model are preserved in discrete terms, except that now the average cost curve may or may not intersect the demand curve. Therefore the *critical mass* x^* is now re-defined as the smallest quantity of users at which willingness to pay exceeds average costs, and we assume $x^* > 1$.

[9] A *game allocation* is a function, $\varphi : X \ \{0,1\}$, i.e. an element of the set $\{0,1\}^X$, a sequence of 0's and 1's indexed by the set X.

[10] i.e. $\varphi^*(x')$ maximizes the payoff to player x', given the values of for all $\varphi^*(x)$ for all $x \in X, x \neq x'$. This is a standard concept of Nash equilibrium in non-cooperative games. For any game allocation φ, let $(\theta, \varphi^0(x))$ be the allocation with values equal to those of φ everywhere except at x, and with $\varphi(x) = \theta$.

[11] i.e. the region of stability E_0 is usage up to critical mass minus one.

[12] Rohles, J., "A theory of interdependent demand for communication services," *The Bell Journal of Economics and Management Science*, Vol. 5 No. 1, pp. 16-37, Spring 1974; and Heal, G.M., "Price and Market Share Dynamics in a Network Industry," Working Paper, Columbia University Graduate School of Business, New York, 1990.

[13] Rohles, J., "A theory of interdependent demand for communication services," *The Bell Journal of Economics and Management Science*, Vol. 5 No. 1, pp. 16-37, Spring 1974.

[14] To simplify computations in the following we assume $x^*=X_{max}$. Obviously when $x^*<X_{max}$ both the decentralized process and the centralized process proceed faster to reach a coalition of average critical mass. This is because when $x^*<X_{max}$, the probability that x^* coins come up simultaneously heads is strictly larger than the probability that all coins is a set of cardinality x^* come up heads simultaneously. Similarly, the formation of α decentralized coalitions of β players each summing up to cardinality x^* within a larger group of cardinality $X_{max}>x^*$ is faster, so that the final result for $x<X_{max}$ is not significantly altered.

[15] Rohles, J., "A theory of interdependent demand for communication services," *The Bell Journal of Economics and Management Science*, Vol. 5 No. 1, pp. 16-37, Spring 1974.

[16] Oren, S. and S. Smith, "Critical Mass and tariff structure in electronic communications markets," *The Bell Journal of Economics*, 12, pp. 467-86, Autumn 1981.

[17] Katz, M. and C. Shapiro, "Network Externalities, Competition and Compatibility," *The American Economic Review*, Vol. 75 No. 3, June 1985, pp. 424-40.

[18] Farrell, J. and G. Saloner, "Standardization, compatibility, and innovation," *Rand Journal of Economics*, Vol. 16, No. 1, Spring 1985.

[19] Heal, G.M., "Price and Market Share Dynamics in a Network Industry," Working Paper, Columbia University Graduate School of Business, New York, 1990.

[20] Oren, S. and S. Smith, "Critical Mass and tariff structure in electronic communications markets," *The Bell Journal of Economics*, 12, pp. 467-86, Autumn 1981.

[21] Katz, M. and C. Shapiro, "Network Externalities, Competition, and Compatibility," *The American Economic Review*, Vol 75. No. 3, pp. 424-40, June 1985.

[22] Farrell, J. and G. Saloner, "Standardization, compatibility, and innovation," *Rand Journal of Economics*, Vol. 16 No. 1, Spring 1985.

[23] Heal, G.M., "Price and Market Share Dynamics in a Network Industry," Working Paper, Columbia University Graduate School of Business, New York, 1990.

[24] Ibid.

[25] Ibid.

[26] Rohlfs, J., "A theory of interdependent demand for a communication services," *The Bell Journal of Economics and Management Science*, Vol. 5 No. 1, pp. 16-37, Spring 1974.

[27] Note that with this definition the critical mass payoffs could be strictly positive; for example this would occur in Diagram 3 if *C(x)* goes between two horizontal steps of *D(x)*.