CENTRO STUDI IN ECONOMIA E FINANZA

CENTRE FOR STUDIES IN ECONOMICS AND FINANCE

WORKING PAPER NO. 61

Network Size, Value and Cycles

Hans W. Gottinger

June 2001

DIPARTIMENTO DI SCIENZE ECONOMICHE - UNIVERSITÀ DEGLI STUDI DI SALERNO
Via Ponte Don Melillo - 84084 FISCIANO (SA)
Tel. 089-96 3167/3168 - Fax 089-96 3169 – e-mail: csef@unisa.it

CSEF **WORKING PAPER NO. 61**

Network Size, Value and Cycles

Hans W. Gottinger

(University of Maastricht, Fondazione CARISAL and CSEF, University of Salerno)

Table of contents

1. Introduction

One of the most striking economic aspects of networks is how they create externalities. Network externalities occur in both the demand and supply of the network. The textbook externality is a supply externality. For example, as a negative byproduct of a factory's production, pollution spews into the air or water. Demand externalities, on the other hand, may exist for non-network goods, but they are not usually considered important enough to merit attention. For example, economists typically do not factor demand externalities into consumers' demand functions. Many models of consumer behaviour assume that the average consumer's demand for potatoes, lettuce, corn, etc., for example, are formed without any reference to how many other people are purchasing these products. Certainly the number of consumers in a given market affects demand and therefore price, but an individual's demand is independent -- it does not depend directly on a product's popularity in most models. Such effects are assumed away as insignificant.

Besides the supply-side economies of scale the demand-side economies of scale are commonly seen in the communications and computer industries among others. For some goods and services, a person's demand depends on the demands of other people, or the number of other people who have purchased the same good may affect a person's demand. For example, the buyer of a telephone or fax machine would have not bought it if there were no one else who had purchased or would have purchased it. When more people have purchased it the more value of a telephone or fax machine the buyer would have obtained. This is a positive network externality based on an 'actual' or 'physical' network. Moreover, for some goods, such as Microsoft Office, the individual demand for that good inherently exists but enormously increases when other people buy the same good. In an actual network, products have very little or no value when alone, they generate value or more value when combined with others (example: fax machine). In a virtual network, hardware/software network products have value even if they exist alone, however, they are more valuable when there are more complementary goods, and also there will be more complementary goods when more people use the products. Application software developers are likely to write for the platform of the operating system that most people favour. Conversely, the operating system that more application software writes on are favoured by more people. The operating system with a larger market share will provide a bigger market for the application programs. At the same time, the availability of a broader array of application programs will reinforce the popularity of an operating system which in turn will make investment in application programs compatible with that operating system more desirable than investment in application programs compatible with other less popular systems. As a result, the operating system with a larger installed base attracts more buyers whereas the small and later entrant with a smaller installed base with equal or even superior quality finds it difficult to compete. As more users are attracted to a network, the size of the network grows and confers greater value to the network users. Network effects directly challenge an important principle of classical economic theory, which posits decreasing (and eventually negative) returns to scale in most markets. Also this theory basically deals with increasing returns problems in case of supply side economies of scale but ignores cases of demand side economies of scale brought about by increasing value of existing users through increased demand, i.e. through network externalities. That is, network markets offer increasing returns over a large portion of the demand curve or even the entire demand curve. Markets with increasing returns imply that bigger is better and consumers deriving more value as the number of users grows. The flip side of this situation in terms of market structure is that the strong grow stronger and the weak become weaker.

Hence, network markets provide potentially fruitful returns to firms that can make their own products as standards in markets or in aftermarkets for complementary goods. This presents the possibility of substantial first-mover advantages: being the first seller in a market may confer an important strategic advantage over later entrants because a first mover's technology may become locked in as a standard (Arthur, 1989, Katz and Shapiro, 1986)That is to say, the first technology that is introduced into the market may gain excess momentum when many early users join in anticipation of other users hopping on the bandwagon at a later date. This strong expectation is critical to network expansion (Choi, 1997). In the end consumers already belonging to an existing network will not likely switch to a new technology, even if it is better (Economides, 1996)

The switching costs associated with transferring to an incompatible but superior technology create 'excess inertia' to consumers. That means consumers will not adopt a new superior technology not only because of the sunk costs they have already put in but also because values from network externalities may be lost if they switch. Network effects, therefore, could stifle innovation.

In a traditional market, where network effects are negligible or non-existent, competition turns primarily upon price, quality and service considerations. In contrast, in those markets in which network effects are significant, competition plays out in other dimensions as well: particularly in strategies to establish, maintain, and control standards for the industry. The computer industry hence suggests that network effects have played an important role in shaping the market structure and the margins on which competition occurs.

Also, increasing returns raise the possibility of leveraging a monopoly power from one market to another. Because users may be reluctant to commit to any given system unless they believe it will be adopted by many others, the 'network owner' may engage in a variety of strategies to discourage potential buyers from buying a smaller network regardless whether or not it is superior. Strategies include expanding the system to include complementary products offering a wide variety of complementary products at very attractive prices or through bundling. At the same time, leveraging is able to raise rivals' economic costs of competing in the marketplace.

For example, in its effort to be adopted as the next generation standard, the owner of one element of a system may enter complementary markets by engaging in alliances as part of a strategy of attracting users to its network. Consequently, rival operating systems need to ensure the provision of substantial complementary products in the market, otherwise very few buyers will try its system. As a result, the follow-on improved or complementary products markets become very difficult.

Strong network effects are therefore themselves barriers to entry, even though it is sometimes unclear whether entry into the market ought to be encouraged. Since the increasing return deters the incentive of new entrants and increases the costs of new entrants. Such a blunting of incentives can occur if the leveraging practice is undertaken, not primarily as part of a vigorous competitive strategy, but in part to decrease the likelihood of competitor entry, so that the dominant firm will continue to be dominant in competition for the next market. This has clearly be shown for the Japanese telecommunications market (Gottinger and Takashima, 2000b) The unlikelihood of success for new entrants will reduce the incentives of other competitors to innovate to the extent that these competitors perceive that the opportunities to profit from their innovations are hindered. All of this is particularly significant because markets in which there is rapid technological progress are often markets in which switching costs are high, in which users find it costly to switch to a new technology that is not fully compatible with the older technology. The result is an increase in entry barriers.

From what follows the definition of a network externality is given by the value of a network created by the number of its nodes. Also, network externalities can exist both for the supply and demand side of the economic equation. And networks can generate negative, positive or no externalities. Network externality networks are those that decrease in value when the number of nodes increases. More 'traditional' network industries fit into this category.

2. *Perspectives on Network Externalities*

We start with a useful distinction suggested by Economides (1996) in his survey of the literature, he divides the work on network externalities into what he calls macro and micro approaches. Macro investigations assume that externalities exist and then attempt to model their consequences. Micro investigations start with market and industry structures in an attempt to derive (theoretically) the source of network externalities. The later category is largely founded on case studies. Three of those are symptomatic. David's(1985) QWERTY study,

Arthur's (1989) model, and the domination of VHS in the videotape recorder market combined, spurred theoretical and empirical interest in network externalities. The gist of David's QWERTY study is that inferior technologies through network externalities may be subject to 'lock-ins'. This might apply to the keyboard QWERTY as well as to the adoption of the VHS against the Betamax standard though with specific technological advantages of Betamax over VHS. In empirical support of network externalities, Gandel (1994) finds that consumers pay a premium for spreadsheets which are compatible with Lotus 1-2-3 (an industry standard for spreadsheets). In other words, consumers are willing to pay for the ability to share spreadsheet information and analysis easily with other computer users. Thus he concludes that there is strong empirical support for the existence of network externalities in the computer spreadsheet market. In another aper, Saloner and Shepard (1990) test for the existence of network externalities in the network of Automated Teller Machines (ATMs), their results support existence.

3. *Hypotheses on Network Externalities*

Perhaps it is not surprising that little quantitative work on network externalities has been done. Many examples of network industries embody cutting-edge technologies, given that theoretical work on network externalities is still relatively new, data collection is fragmentary and common data sets upon which to test theories are severely limited. One particular important question emerging on network externalities is the functional relationship between the size of a network (its number of nodes) and the network's value.

Three key assumptions about the relationship between network size and network value underlie most analyses of network externalities and their effects. They relate to linear, logarithmic and exponential assumptions.

The linear assumption postulates that, as networks grow, the marginal value of new nodes is constant. The logarithmic assumption postulates that, as a network grows, the marginal value of new nodes diminishes. Network externalities at the limit in this formulation must be either negative, zero or of inconsequential magnitude in comparison to quantity effects on prices. In contrast, Katz and Shapiro (1985) make their assumptions explicit: network externalities are positive but diminish with development, at the limit they are zero.In any case, network effects diminish in importance in these models as a network grows. The third key assumption about the relationship between network size and value is the exponential assumption which in the popular business and technology press has been named 'Metcalfe's Law'. It embodies the idea of positive network externalities whose marginal value increases with network size. Robert Metcalfe (1995) states the 'law' in this way: "In a network of N users, each sees a value proportional to the N-1 others, so the total value of the network grows as N(N-1), or as N squared for large N". The validity of Metcalfe's Law is crucial to the 'increasing returns' debate on the New Economy, facilitated by the aggregation of positive network externalities in high tech industries. One could also consider a mixture of hypotheses such as a combination of Metcalfe's Law and the logarithmic assumption, that is early additions to the network add exponentially to the value of a network, yet later additions diminish in their marginal value. The result looks like an S curve, as illustrated.(Figures 1 and 2) It is based on the idea that early additions to a network are extremely valuable, but at some point 'network saturation' should take place and marginal value should fall.

Figure 1 : Three Assumption of the Network Size/ Value Relationship

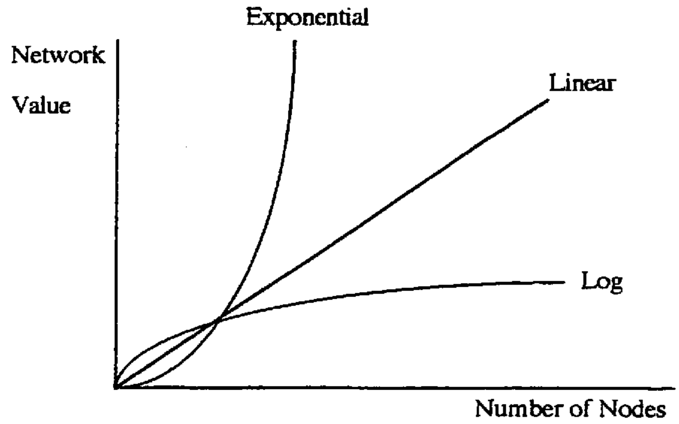
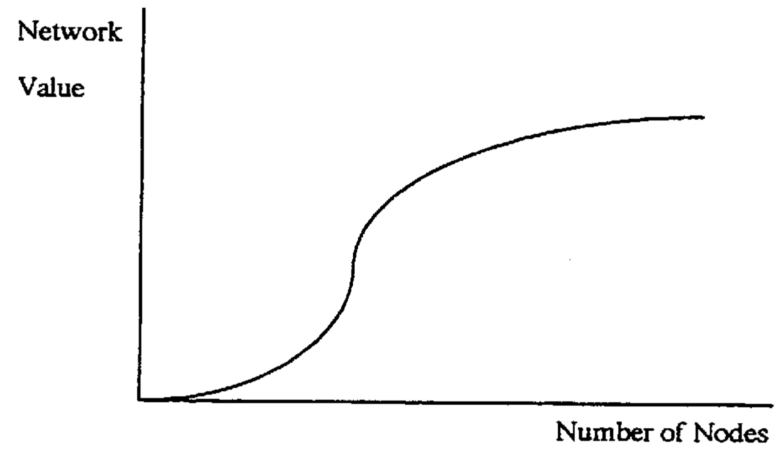


Figure 2 : The S-Hypothesis



In summary, the industry and hence aggregate (growth) benefits can be classified as follows:

(i) industries that show an exponential growth (through strong complementarity)

(ii) industries that show linear growth (additive benefits)

(iii) industries that show a log relationship (stable benefits)

The mixtures of those economies create the features of the new economy. Such an economy is not immune to economic cycles, but to the extent that the network economy snowballs in an upswing period, by the same token it might also along the supply chain contract faster in a downswing period but with a better chance to stabilize quicker.

4. Technology Adoption, Network Industries and Network Effects

We look at the main hypotheses as how they are likely to affect the adoption process of particular network industries. The linear hypothesis is the assumption of Arthur's (1989) model subject to simulation. Given a very large number of trials, technology adoption leads (almost surely) to lock-ins. Given two technologies, A and B, further R and S agents that make adoption decisions, respectively, in Arthur's model each trial represents a random walk of an ever increasing number of R and S agent decisions. As the number of trials increases, with symmetries in both technologies A and B, the split between A and B adoptions approach fifty-fifty. That is, either one of them will be adopted, and non-adoption will be most unlikely.

In Arthur's analytical model, as the number of iteration goes to infinity, the possibility of non-adoption disappears.

Correspondingly, the average adoption time until lock-in will increase with decreasing probability (of non- adoption), in conformity with the linear hypothesis, in other words, more agents become (linearly) more convinced to adopt either way. This suggests that the network effect leaves only a neutral impact on the innovation process. Against this bench mark, when the value of network size grows logarithmically in relation to its size, the average time until lock-in occurs is extended. What appears surprising is how much the logarithmic assumption delays lock-in. That is, the logarithmic specification creates less growth prospects and greater instability by delaying (or preventing) adoption from occurring. In contrast to the logarithmic hypothesis, the exponential assumption shortens the average time until adoption occurs. The average adoption is affected just as drastically by the exponential assumption as by the logarithmic one. With the exponential assumption, however, the average adoption occurs much earlier than in the baseline case. No wonder, that on an aggregate scale across network industries, it is this network effect that lends support to 'increasing returns' by the proponents of the New Economy. It can even be reinforced by speed of transactions, for example, enabled through large scale broadband internet technologies. This would support a scenario of a sustained realization of an exponential assumption as even more likely. If it can be established that the Internet triggers a technology adoption process in the form of a large and broad wave ('tsunami') across key industries, sectors, regions and countries , then increasing returns will generate exceptional growth rates for many years to come. For this to happen there should be a critical mass of network industries being established in an economy. Then an innovation driven network economy feeds on itself with endogeneous growth. It remains to be determined, empirically, which mix of sectors, with network effects with exponential, linear and logarithmic relationships will have a sustained endogeneous growth cycle.

From a slightly different perspective, it is interesting to note that the logarithmic assumption creates instability in the models. Metcalfe's law, on the other hand, which leads to immediate adoption, creating a dynamics of its own, would prevent many contemporary models from reaching equilibrium.

5. Networked Industrial Organization

The development of the Internet and its use for business transactions, B 2 B or B 2 C., would make a good subject for the analysis of business cycles for several reasons. First, the Internet might constitute a formal mapping of business transactions for a networked economy, or major parts of their industries. Second, the Internet itself can be modelled and designed as a set of transactions among various agents, consumers, network suppliers and services, that reflect the size and complexity of a distributed computing system (Gottinger, 2000a). That is, the Internet provides prototypical examples of a positive network externality industry. Unfortunatly, the Internet is 'too recent' to make it an eligible candidate for a statistical study on the value of a network. Instead we use the US telecommunications network over the past 50 years. There are some intuitive reasons. In many industry studies, where positive network externalities are defined, Economides (1996) and the Economist (1995), telecommunications is named as the signature network industry. The telecommunications industry is somehow described as a 'typical' network industry that makes it a logical place to begin a search for empirical evidence. Due to its structural similarity with other network industries like railroads, airlines, and the Internet, conclusions reached about network externalities in the communications system are arguably applicable to all of the rest.

In correspondence to the classification (i) to (iii) in Sect. 2 we would venture the hypothesis that (i) exponential growth would likely be associated with an emerging, broadly based advanced technological, strongly growing network industry, (ii) a linear relationship would be tantamount to a maturing, structurally stable industry, while (iii) a logarithmic shape would go with a technologically mature, well established industry.

The weighting of factors (i) to (iii) would characterize the scope and degree of a new economy, the higher the share of (i) and possibly (ii) the stronger the scope of a New Economy though a sizable share of (i) to (iii) would form the basis of a New Economy.

We conjecture that the higher (i) and (ii) in a sizable share of (i) to (iii) the stronger the cyclicity of the economy and the higher the volatility of movements..

6. Regression Specification

Regression analysis can test the three hypotheses about the relationship between network size and network value. Determining "network value" is an obvious difficulty for empirical research done on network externalities. Gandal (1994) addresses the issue of network value by analysing prices: evidence of a price premium for products with a large network indicates that consumers value network externalities positively. The price approach, however, would be flawed if applied to the US telephone network due to its heavy regulation from the 1920s through the 1982 divestiture of AT&T. An alternative approach in such a regulated environment is to observe netwok usage as a proxy for network value, with greater network usage indicating greater network value.

Based upon conventional economic theory and ideas about network externalities, there are six variables which are likely to determine use of the US telephone network and which should therefore be considered in a regression analysis:

1) network size (number of nodes),
2) price of using the network;
3) the price of substitute services;
4) the population ;
5) income; and
6) investment in the telephone network. Each of these variables can bequantified either through data that US telephone companies report or through standard statistical measures of the US population and economy: network size is recorded in the number of telephones active in a given year; average prices for the telephone and its major historical substitute (the telegraph or later the fax)) can be derived from company operating revenues; telephone system capital investment is tracked: and population and income are standard data kept by the US government. Moreover, network usage can be quanttified as the number of telephone conversations per year.

Ad hoc conjecture, based upon standard economic logic, argues for the following predicted signs. Average price of a telephone call should be negatively correlated with network use. Population, income, capital investment, and average price of telephone substitutes should be positively correlated with telephone network use. These predictions are summarized in Table 1:

TABLE 1 : REGRESSION VARIABLES AND PREDICTED SIGNS

Variable	*Proxy*	*Data*	*Expected Sign*
Network Value (dependent)	Network Usage	Number of phone conversations/year	
Network size		Number of "telephone sets"	Positive*
Price of network use		Average cost of phone call	Negative
Substitute goods' price		Average cost of telegram	Positive
Income		Real GDP	Positive
Telephone system investment		Capital investment/year	Positive
Regulatory change	Dummy		

* See text for more detail on the expected sign of the network size variable

The sign of the relationship between network usage, of course, is the primary focus. Two methods for specifying and testing the hypotheses regarding network size and usage could be used. The first method specifies a separate model for each hypothesis and then compares the fit of each model to the data. The second method uses a model general enough that any of the three hypotheses could be valid, and then uses the estimated beta coefficients to determine the most likely specification. We use the latter method because the model specifications necessary for the first method would be difficult to compare on the basis of fit (specifically, using a logarithmic transformation makes comparison to untransformed data problematic).

The second more general method requires that the variable's specification must be flexible enough to produce any of the functional forms of the hypotheses. Thus we use a polynomial equation to evaluate the relationship between network size and usage. Given a polynomial of sufficient degree, any functional form can be reproduced. Equation (1) shows the general form of the regression model, where y is the dependent variable and the b's represent the beta coefficients.

$$y = a + b_1 \text{size} + b_2 \text{size}^2 + b_3 \; size^3 + b_4 \, Pphone + b_5 \, Ptel + b_6 \, GDP + b_7 \, Dummy + \ldots \quad (1)$$

Population does not appear as a separate variable in this formulation because all non-price variables have been transformed to per capita measures.

Estimation of the beta coefficients in equation (1) should yield evidence as to which of the hypotheses most accurately represents the ralationship and value in the telephone network. For example, if b_1 and significant while b_2 and b_3 are not sinificantly different from zero, this evidence would support the linear hypothesis (ii). In contrast, if b_2 were the only positive and significant coefficient, this would lend support to the assumptions of Metcalfe's law.

The logarithmic and S-shape hypotheses could be simulated through several combinations of positive and negative coefficients among the polynomial terms, but a mixture of signs coupled with significant coefficients for the higher order polynomial terms could argue for either the logarithmic or S-shape hypotheses. In any case, examining the signs of the variables will serve as both a check for consistency of the model specification against theoretical expectations and as an argument for which hypothesis best fits the data.

Econometric Issues. One potential problem of working simultaneously with a dependent variable like network usage and the independent variables of network size, income, and network investment is that these variables are likely to trend upward together. The similar growth patterns of these four variables can then make it awkward to disentangle the effects of each independent variable on variation in the dependent variable. Such a situation may also tend toward the presence of heteroskedacity and serial correlation. Three measures are taken here to guard against the potential problems of working with these time series. First, we add a lagged dependent variable to the independent terms to remove significant time trends. Second, we perform the regression on both the untransformed dependent variable and ist natural log to check for significant potential problems with residuals. And third, we use the Durbin-Watson and Cook-Weisberg tests to check for serial correlation and heteroskedasticity.

7. Regression Results

The reggression results overall support the linear assumption of (ii) although there is some evidence that the exponential assumption of (i) is valid for a group of new network industries.

During the course of the analysis, we tested several versions of the regression model. These versions varied primarily in the number of network size polynomial terms (Model with up to the fifth order polynomials were tested), the inclusion of a lagged dependent variable, and in the inclusion of interaction terms. The need to correct for a strong time trend and the existence of serial correlation and /or heteroskedacity in early model specifications led to the inclusion of a lagged depenedent variable and the use of the Cochrane-Orcutt procedure. Also the investment variable was dropped from the model because of multicollinearity problems as it correlated nearly perfectly with the variable measuring network size.

Table 2 presents the full regression equation with estimated coefficients, standard errors of the coefficients, and t-statistics.

The full regression shows mixed results with two outstanding points. First, the linear, squared, and cubic terms of network size are significant at the respective confidence levels of one, five, and ten percent.

TABLE 2: FULL REGRESSION RESULTS

Predictor	*Variable*	*Coefficient*	*Standard Error*	*t-statistic*
Doonvlag	Lagged Dependent	0893283	0986104	0.906
Netsize (NS)	Number of telephone Sets	1550.075	251.3466	6.167***
NS2	NetSize squared	-1291.108	678.0486	-1.904**
NS3	Netsize cubed	1471.154	727.8555	2.021*
Tgrafh	Average real price of telegram	10.35254	9.395631	1.102
GNPpc	Per capita real GNP	-0037915	0062658	-0.605
Avgcallp	Average real price of telephone call	-91.92725	149.9191	-0.613
Monop	Structural change dummy	-2.168961	6.681732	-0.325
Inter	Constant	3.082576	22.13733	0.139
Rho	Cochrane-Orcutt correction	0.8036	0.0732	10.973***
		No. of obs.=67		
		R-sq=98.7%	AdjR-sq=98.6%	

*** Denotes variable significant at the 1% level

** Denotes variable significant at the 5% level

* Denotes variable significant at the 10% level

This indicates some nonlinearity in the relationship between network size and value. Second, the other explanatory variables in this specification covering income and substitution effects are not significant.

These regression results are not especially convincing, given the lack of significance for all price variables, but this regression serves as a useful stepping stone in the econometric analysis. The important piece of information to glean from this full regression is the estimated relationship between network size and value derived from the network size coefficients and the data. Forecasting a doubling in the density of telephones per capita in the US using the estimated coefficients displays an upward trend of a weak exponential relationship. Because the upward trend appears only in forecast, though, the evidence for the exponential hypothesis should be considered weak on average.. Because the three network size terms essentially estimate a linear relationship between the number of telephones per capita and the number of domestic conversations per capita per year, a logical modification to the full regression specification given above is to eliminate the network size squared and cubic terms. Respecifying the regression in this manner results in an increase in the t-statistics for both the price variables, though the telephone call price variable can still not be assumed to be non-zero. The full results of this final regression appear in Table 3.

TABLE 3: FINAL REGRESSION RESULTS

Predictor	*Variable*	*Coefficient*	*Standard Error*	*t-statistic*
DconvLag	Lagged Dependent	- 0969381	-0983045	0.986
Netsize (NS)	Number of telefhone sets	1216.536	140.8213	8.639***
Tgraph	Avarage real price of telegram	16.05634	5.843993	2.747***
CNPpc	Per capita real GNP	-0036921	0063857	-0.578
Avgcallp	Avarage real price of telephone call	-151.5678	140.7213	-1.077
Monop	Structural Dummy	-1.440889	6.617304	-0.218
Constant		18.70185	12.24442	1.527
Rho	Cochrane-orcutt procedural variable	0.8442	0.0626	13.492***
		No. Of obs.=67		
		R-sq=98.1%	AdjRsq=97.9%	

*** Denotes variable significant at the 1% level

** Denotes variable significant at the 5% level

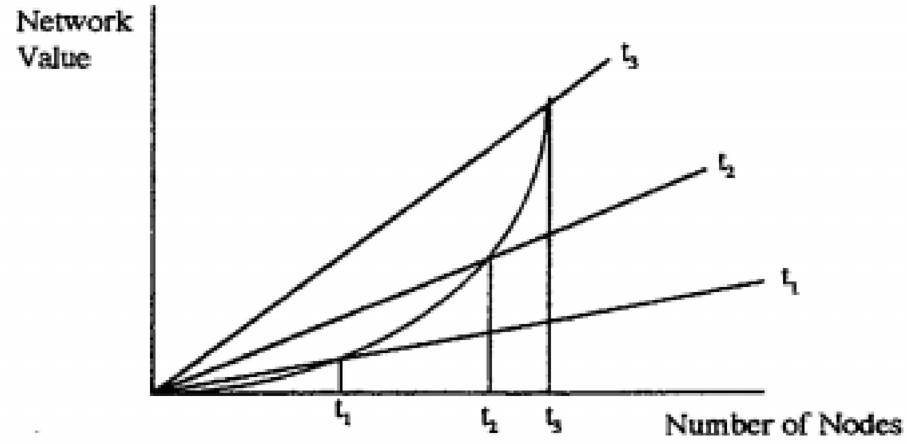
* Denotes variable significant at the 10% level

Two clear concerns in a model with significant time trends and a high adjusted R-squared measurement are heteroskedasticity and autocorrelation. Neither of the model specifications, however, displays signs of these problems. Fitted values versus residuals plots show no apparent problems with the residuals. Moreover, the Cook-Weisberg test for heteroskedasticity cannot reject the null hypothesis of costant variance and the Cochrane-Orcutt procedure for correcting autocorrelation has been employed. The subsequent Durbin – Watson h-statistic indicates no autocorrelation at a one percent level of significance for each of the specifications.

8. Effects of Technological Change

In the present case, one difficulty to observe variables that may affect the relationship between network size and network value and thus create problems in the results is technoloical change. Technological improvements, through better quality, value added features, high speed access and the like, could cause an upward shift in the relationship between network size and value. In other words, given a certain network size, technological improvements alone could lead to greater network usage. Fig. 3 shows the potential problems that shifts in the true relationship can cause for estimations. The straight lines t_1, t_2 , t_3 show the true relationship as it shifts up; the curved line illustrates how sampling at different times can estimate an upward -curving relationship when it is really linear.

Figure 3 : Potential Distortion over Time



9. Conclusions

For various key technological sectors or industries we can identify a strong network economy with an exponential network value to size relation, as compared to one with a linear or logarithmic relationship. Such an economy is characterized by rapid growth and innovation trajectories reinforced by expectations on queueing constraints. As the dynamics unfolds, lack of coordination through competriutive pressures creates excess capacities which with a slackening demand supported by diminiished expectations leads to swift, drastic and steep decline avalanching through the supply chain network. An exponential type network economy would trigger significantly more volatile business cycles than we would see in any network economies, or even in the rest of the old economy. The level or extent to which a strong network economy affects the overall economy, such as the US economy recently, even less so physically than in perception, is tantamount to a serial negative queuing effect which is not supply but demand queuing.

References

Arthur, W.,1989, Competing technologies, increasing returns, and lock-in by historical events, *Economic Journal* 99, 116-131

Choi, J.P., 1994, Irreversible choice of uncertain technologies with network externalities, *Rand Journal of Economics* 25 (3), 382-401

David, P., 1985, CLIO and the economics of QWERTY, *American Economic Review*, PP 75, 332-337

Economides, N.,1996, The economics of networks, *International Journal of Industrial Organization* 14, 673-699

The Economist, 1995, A Survey of Global Telecommunications, September 1995

Gandal, N., 1994, Hedonic price indexes for spreadsheets and an empirical test for network externalities, *Rand Journal of Economics* 25(1), 160-170

Gottinger, H.W., 2000a, Network economies for the Internet, *Netnomics* (to appear)

Gottinger, H.W. and M. Takashima,2000b, Japanese telecommunications and NTT Corporation: a case in deregulation, *Inter. Jour. of Management and Decision Making* 1 (1),68–102

Katz, M. and C. Shapiro,1986, Technology adoption in the presence of network externalities, *Journal of Political Economy* 94, 822-841

Metcalfe, R., 1995, Metcalfe's Law: a network becomes more valuable as it reaches more users, *InfoWorld* 17 (40), 53

Studenmund,A.H.,1992, *Using Econometrics:A Practical Guide*, 2nd ed., Harper Collins, New York.

Statistical Appendix: *Data Sources*[*]

Data for regressions performed come partly from Historical Statistics of the United States (HSUS), from l950 to 1970, for the remaining period up to 2000 annual FCC reports have been used. The time series have been constructed from various data sets over long periods of time, the notes on these data are extensive.

i. *Elimination of duplicate figures*. In HSUS two separate figures exist for a number of years (see for example, data on the number of telephone instruments, the number of calls made, plant, revenue, and cost information). Usually, one figure is derived from FCC reports, and the other is culled from census data. These data do not always match. The yearly time series data were maintained, and the census material eliminated from the data set.

ii. *Creation of the dependent variable.* Time series in HSUS give the average daily number of telephone calls in the U.S. from 1950 to 1970. To eliminate international data that exist in these time series but should not appear in the dependent variable time series, the new figures were obtained by subtracting overseas phone calls from total conversations per year.

iii. *Per capita transformation.* The dependent variable (domestic conversations), the lagged dependent variable, and the independent variables network size, network size squared, network size cubed, and gross national product were transformed to per capita measurements using census figures.

iv. *Creation of Price Variables.* To create a figure for the average cost of telephone conversation, total telephone industry operating revenue was divided by the total number of conversations. Similarly, the average price of a telegram were calculated through the operating revenues of the entire domestic telegraph industry from 1950 to 1980. These operating revenues were divided by the total number of telegrams handled to arrive at the average figure. All price data and the time series for GNP were converted to real figures using historical consumer price index information. The base year is 1967.

v. *Full Regression Specification*

We present the full regression results for the three regression specifications discussed in the text.

[*] For the regression calculations I am indebted to Jan van Straaten, Univ. of Maastricht.

Full Regression Specification

(Cochrane-Orcutt regression)

Iteration 0:	rho =	0.0000
Iteration l:	rho =	0.3183
Iteration 2:	rho =	0.5303
Iteration 3:	rho =	0.6652
Iteration 4:	rho =	0.7374
Iteration 5:	rho =	0.7748
Iteration *6:*	rho =	0.7922
Iteration 7:	rho =	0.7994
Iteration 8:	rho =	0.8022
Iteration 9:	rho =	0.8032

source	SS	df	MS
Model	156259.631	8	19532.4538
Residual	2010.38066	58	34.6617355
Total	158270.011	66	2398.03047

Number of obs		50
F (8, 58)	=	563.52
prob > -F	=	0.0000
R-squared	=	0.9873
Adj R-squared	=	0.9855
Root MSE	=	5.8874

DomConv	Coef.	Std. Err.	t		[95% Conf.interval]	
DconvLag	-0893283	.0986104	0.906		-.1080619	.2867185
NS	1550.075	251.3466	6.167	0.000	1046.95	2053.2
NS2	-1291.108	678.0486	-1.904	0.062	-2648.369	66.15401
NS3	1471.154	727.8555	2.021	0.048	14.19284	2928.115
Tgraph	10.35254	9.395631	1.102	0.275	-8.45486	29.15994
GNPpc	-.0037915	.0062658	*-0.605*	0.547	-.0163339	.0087509
AvgceLUp	-91.92725	149.9191	-0.613	0.542	-392.023	208,1685
Mmop	-2.168961	6.681732	-0.325	0.747	-15.5439	11.20598
-inter	3.082576	22.13733	0.139	*0.890*	-41.2301	47.39525
rho	0.8036	0.0732	10.973	0.000	0.6573	0.9498

Durbin-Watson statistic (original)	1.360806
Durbin-Watson statistic (transformed)	1.923850

Final Regression Specification

(Cochrane-Orcutt regression)

Iteration 0:	rho =	0.0000
Iteration l:	rho =	0.4283
Iteration 2:	rho =	0.5829
Iteration 3:	rho =	0.6910
Iteration 4:	rho =	0.7650
Iteration 5:	rho =	0.8077
Iteration *6:*	rho =	0.8288
Iteration 7:	rho =	0.8381
Iteration 8:	rho =	0.8420
Iteration 9:	rho =	0.8436

source	SS	df	MS
Model	108242.777	6	18040.4629
Residual	214444.31715	60	35.7386192
Total	110387.094	66	1672.53173

Number of obs		50
F (6, 60)	=	504.79
prob > -F	=	0.0000
R-squared	=	0.9806
Adj R-squared	=	0.9786
Root MSE	=	5.9782

DomConv	Coef.	Std. Err.	t		[95% Conf.interval]	
DconvLag	.0969381	.0983045	0.986	0000	-.0997001	.2935763
NS	1216.536	140.8213	8.639	0.000	934.8516	1498.221
	-		-	0.062	-	
				0.048		
Tgraph	16.05634	5.843993	2.747	0.275	4.366612	27.74606
GNPpc	-.0036921	.0063857	*-0.5785*	0.547	-.0164654	.0090813
Avgcallp	-151.5678	140.7313	-1.077	0.542	-433.0523	129.9168
Monop	-1.440889	6.617304	-0.218	0.747	-14.6774	11.79569
-inter	0.132	-5.790629	43.19433	*0.890*		
rho	0.8442	0.0626	13.492	0.000	0.7193	0.9691

Durbin-Watson statistic (original)	1.142124
Durbin-Watson statistic (transformed)	2.002030

Durbin-Watson Statistics

The transformed Durbin-Watson statistics for the two regression specifications are:

(full) 1.923850, and (final) 2.002030. Because a lagged dependent variable is used in the regressions, the Durbin-Watson statistic cannot be used to judge the presence of autocorrelation (Studenmund, 1992), rather, the D-WQ statistic must be transformed into Durbin's h-statistic. Durbin's h-statistic for the two specifications are, respectively, 0.38611546 and –0.0103251, these statistics do not allow a rejection of the null hypothesis of no serial correlation at the ten percent confidence level.

Cook-Weisberg Tests

The Cook-Weisberg tests for the models do not support the finding of heteroskedacity

Full Specification

Cook-Weisberg test for heteroskedacity using fitted values of DomConv

Ho: Constant Variance
chi2(1) = 0.35
Prob > chi2 = 0.5555

Final Specification

Cook-Weisberg test for heteroskedasticity using fitted values of DomConv

Ho: Constant Variance
chi2(1) = 0.53
Prob > chi2 = 0.4655