PREDICTING ELECTRONIC COMMERCE GROWTH: AN INTEGRATION OF DIFFUSION AND NEURAL NETWORK MODELS

Somnath Mukhopadhyay[1]
Information and Decision Sciences,
The University of Texas at El Paso,
El Paso, Texas 79968-0544.
smukhopadhyay@utep.edu

Subhashish Samaddar
Department of Managerial Sciences,
Georgia State University,
Atlanta, GA, 30303.
s-samaddar@gsu.edu

Satish Nargundkar
Department of Managerial Sciences,
Georgia State University,
Atlanta, GA 30303.
snargundkar@gsu.edu

ABSTRACT

There is a growing recognition that e-market planners and various planning agencies in Information Technology sectors have a significant interest in measuring and forecasting the growth of e-commerce. The difficulties lie in finding a forecasting model that can incorporate both internal and external influences on diffusion, as well as an acceptable measure for e-commerce growth. This study uses models based on the knowledge of traditional diffusion theories as well as artificial neural networks. Additionally, it integrates the two into a hybrid model in order to study e-commerce growth. A count of dot-com hosts is used as a reliable measure of e-commerce growth in all the models. Our study demonstrates that a simple Neural Network model, if properly calibrated, can create a very flexible response function to forecast e-commerce diffusion growth. The neural network model successfully modeled both the internal and external influences in the data, while the traditional formulations could only model the internal influences. The predictive validation of the results was enhanced by replicating the comparisons on simulated data with various degrees of external influence. The study suggests that when external influences are present, the neural network model will be superior to the best traditional diffusion model.

Keywords: E-commerce, Dotcom, Forecasting, Neural Network, Diffusion models, and E-market Planning

1. Introduction

Studying the diffusion of e-commerce is extremely important for both government and business investors and policymakers for effective planning [Press 1997; Yao 2004]. However, industry and academic researchers found that measuring, forecasting and tracking the global diffusion of e-commerce faces two hurdles. The first problem is one of appropriately modeling the diffusion, both to understand the phenomenon and to forecast the diffusion for planning purposes. The process of innovation diffusion has been extensively researched [Rogers 1983], and several traditional diffusion models have been used to explain and forecast the phenomenon. Significant research in the past has used such models for the explanation and prediction of the diffusion of different technological innovations – e.g., the Bitnet [Gurbaxani 1990] organizational forms [Mahajan et al. 1998], corporate governance mechanisms [Venkatraman et al. 1994], the Internet [Rai et al. 1998], web-based shopping system [Changsu & Galliers 2004] and so on. These models

[1] This research was partially supported by a grant from the University Research Institute at the University of Texas at El Paso in 2004 and a summer research grant from the college of business administration at the University of Texas at El Paso in 2005 to the first author.

are therefore a logical first choice in any attempt to understand and forecast e-commerce growth, but they do have some limitations.

The second problem is the difficulty in measuring the growth of diffusion of innovation. For some technological innovations, measurement of diffusion can be obvious. For instance, measuring the diffusion of cell phones is simply a matter of tracking sales of cell phones or subscriptions. Tracking diffusion of e-commerce is much more difficult. One obvious measure, sales dollars, is much less clear, since the portion of sales of each firm attributed to the Internet is not precisely tracked by most firms. In Section 2 we elaborate on the issue of e-commerce and the measurement of its growth.

The key idea of traditional diffusion modeling (discussed in greater detail in section 3) is to assume that there are a fixed number of potential adopters of new innovations. Therefore, this adoption process targets a decreasing number of adopters as time goes by. The growth rate of adoption can be constant [Fourt & Woodlock 1960]. [Mansfield 1961] proposes that the diffusion process follows a simple logistic curve (s-shaped) over time through imitation. [Bass 1969] suggests two main factors that are responsible for the diffusion growth process: *imitation (or contagion)* and *innovation*. [Mahajan & Muller 1979] later called them *internal* and *external influences*. Despite the popularity of the Bass model in explaining the diffusion of innovation of new products, research identified its drawbacks in forecasting growth in the near future. [Heeler & Hustad 1980] point out its instability when data is limited, when there are differences in the environment in which diffusion occurs, as well as a systematic underreporting of estimated time to attain total number of first purchase sales. [Golder & Tellis 1997] suggest that the Bass model fails to predict takeoff in the growth curve properly. [van den Bulte & Lilien 1997] point out that the estimation of unknown ceilings of total number of adopters is often closer to the number of adopters in the last observation period than it is to reality. This drawback may cause it to seriously underestimate the market potential of new technology.

Diffusion models, based on the S-curve or its variations – all using a predetermined relationship, are unable to account adequately for external factors (government interference, for instance) that may cause sudden changes to the diffusion rate. The modeling techniques are inflexible, in that the approach starts with a parametric function selected a priori. The calibration process attempts to fit the data by adjusting the values of the shape parameters. If the diffusion growth were to stop for a while due to an external factor, the S-curve starts to flatten out without any scope of regenerating the growth back into the curve.

Making the assumption that e-commerce growth has been (and may continue to be) affected by external factors like government involvement, security issues, etc., it makes sense to try a more flexible approach to modeling its diffusion. Artificial Neural Networks (ANN, or simply NN) is a logical choice for such modeling, since it has two key advantages over the traditional methods. Neural Networks are flexible in looking for signals in data as opposed to using a predetermined one-dimensional relationship, and it does not require any assumptions about the distributions of data. This modeling technique is sometimes called a connectionist approach because of neural-like connections in the network. [Rumelhart et al. 1986] provide a summary of specific types of connectionist networks. [Hecht-Neilsen 1990] reviews a mathematical treatment of such networks. The applications of NN to a variety of problems are well documented [Elman & Zipser 1987; Sejnowski & Rosenberg 1987; Tam & Kiang 1992]. Research showed that flexibility and generalization are the two most powerful aspects of NN modeling [Sarle 1995; Wieland & Leighton 1988] if NN models are designed properly. NN models are capable of modeling complex patterns in data, and they can be combined with other models to further improve prediction performance. [Barnden 1995] discusses the relative merits of NN modeling over traditional forecasting methods.

NN models do not always generalize for many applications when used for prediction in extrapolation unless these models are designed properly to fit the application phenomenon [Roy & Mukhopadhyay 1997]. This is because the gradient search technique may find a local minimum in the least mean squared cost function instead of the global minimum [Lippmann 1987]. Another difficulty with the NN training algorithm is that in many cases the amount of training data required for convergence is large. For many practical applications, such as the current study where NN models are used to learn the growth pattern of e-commerce, there needs to exist a mathematical function relating correct outputs to inputs with the desired degree of accuracy [NN FAQ 2004] from a limited number of available training data. In order to be a competitive alternative to the traditional models, NN models must achieve at least the same degree of accuracy as the traditional ones. Even with proper specification of the NN models, their performance, though promising, has not been shown to be unequivocally superior [Adya & Collopy 1998; Sharda & Patil 1992]. The challenge for this research is to overcome these difficulties associated with NN models and show that the models generalize well for predicting technological growth in general and e-commerce in particular.

Our first research objective, then, is to assess whether an NN-based integrated approach is a superior alternative to traditional approaches for modeling and forecasting innovation diffusion in general and e-commerce in particular. Superior forecasting alone has obvious practical benefits, and if, through simulation one can show that the combined model would generally perform better in forecasting innovation diffusion, that would be a significant contribution to research. In addition, the contribution to practice comes from increased confidence in forecasts that lead to better planning and policies for both business and government when dealing with the diffusion of innovation. This research intends to achieve the first objective through detailed empirical analysis of real e-commerce growth data as well as simulated data to determine if the results might generalize further.

The second objective of this study is to gain insight from theory into why the NN-based models might be superior (or inferior) to the traditional diffusion models. Specifically, to examine what can be said about the theory of innovation diffusion with respect to the factors influencing the diffusion being investigated. This objective can be achieved by examining existing theory and discussing different diffusion dynamics such as internal influence and external influence on the growth process. Our expectation from the literature is that the combination of internal and external influences is more likely to be modeled better by the hybrid model (diffusion model and NN) for e-commerce growth compared to its traditional counterparts due to the flexibility of NN models, and hence will lead to better forecasts overall when both influences are present. In other words, the NN models could offer a more complete mathematical representation of the phenomenon of innovation diffusion than traditional S-curves.

Accomplishing the above objectives will result in significant contributions for modeling e-commerce growth as well as the growth of other similar technological innovations. Most importantly, it would demonstrate that where an external influence is present, an NN-based model is better than traditional models, and conversely, if NN-based models do not perform differently from traditional ones, one may conclude that the influences on diffusion are purely internal in nature. The results of this study provide alternative views of modeling the growth of innovative products in general and e-commerce growth in particular. It would provide a way to handle external and internal influences together in the same model. Success in modeling and capturing the general impact of external factors - such as governmental investments; tax policies (e.g., no sales tax on Internet transactions); market movements that inject or take away capital infusion to organizations; interoperability and IT standards – would make significant contribution to the field of predicting e-commerce growth.

The rest of the paper is presented as follows. In *Section 2* we discuss e-commerce and its measurement. In *Section 3* we examine the diffusion phenomenon and its modeling alternatives in greater detail. *Section 4* explains the research methodology, while *Section 5* summarizes the results. In order to further validate the results, we performed simulations that are explained in *Section 6*. In *Section 7* we conclude with a discussion of the implications of our findings and provide directions for future research.

2. E-Commerce Growth and its Measurement

The measurement of e-commerce growth is complicated by the difficulty in isolating e-commerce activities of firms from their other activities. Also, e-commerce is but a part of the more complex Internet economy and it may be worthwhile to examine this briefly. A study by [Cisco & the University of Texas 1999] that collected inputs from some 3000 organizations, classifies the Internet economy into four significant segments: (1) *Infrastructure*, (2) *Applications*, (3) *Intermediaries* and (4) *e-Commerce*. Each of the four segments normatively consists of companies that play a critical role in creating and enabling the Internet economy - creating the Internet backbone, providing applications, and Internet commerce and services.

The *infrastructure* segment consists of companies that have products and services that help build the IP-based network infrastructure [Wang et al. 2004]. Such companies include Internet backbone providers such as Qwest, MCI WorldCom to network hardware and software companies such as Cisco, Lucent and 3Com. The second segment consists of Internet *Applications*. Companies in this segment build and provide products and services that build upon the first segment of Internet infrastructure and make it possible to perform business activities online over the IP network. These companies include Internet commerce application providers such as Netscape, Microsoft, etc., and search-engine vendors like Google. *Intermediaries*, the third segment of the Internet economy, facilitate the meeting and interaction of buyers and sellers over the Internet by creating Internet real estate where identification, negotiations, and market making activities are carried out for a fee. This segment includes market makers such as online brokerages like eBay, and as well as content aggregators like Yahoo. The fourth and the final segment in the Internet economy is *e-commerce*, which includes all those involved in sales of products and services to consumers or businesses over the Internet. The players in this segment exist to add economic value to their customers' (both businesses and consumers) otherwise traditional value chain. The interdependence of the Internet segments is illustrated in Figure 1.

Since merchants and infomediaries can conduct their economic activities only when requisitely enabled and supported by the infrastructure and application segments of the Internet economy, it follows that investment in the infrastructure influences the level of e-commerce activity. Along this reasoning, [Gurbaxani 1990] and [Rai, et al. 1998] used number of hosts as a measure of related activity – the former studied the Bitnet, and the latter, the diffusion of the Internet. Following their work, we therefore decided on the number of dotcom hosts as a measure for e-commerce growth in our study. We define dot-com as any website intended for business use to sell any kind of product or service. In the popular media, the term (or the plural form, dotcoms) may mean web-based businesses, referring mostly to ones that failed or suffered cutbacks during the period of March, 2000 to October, 2002.

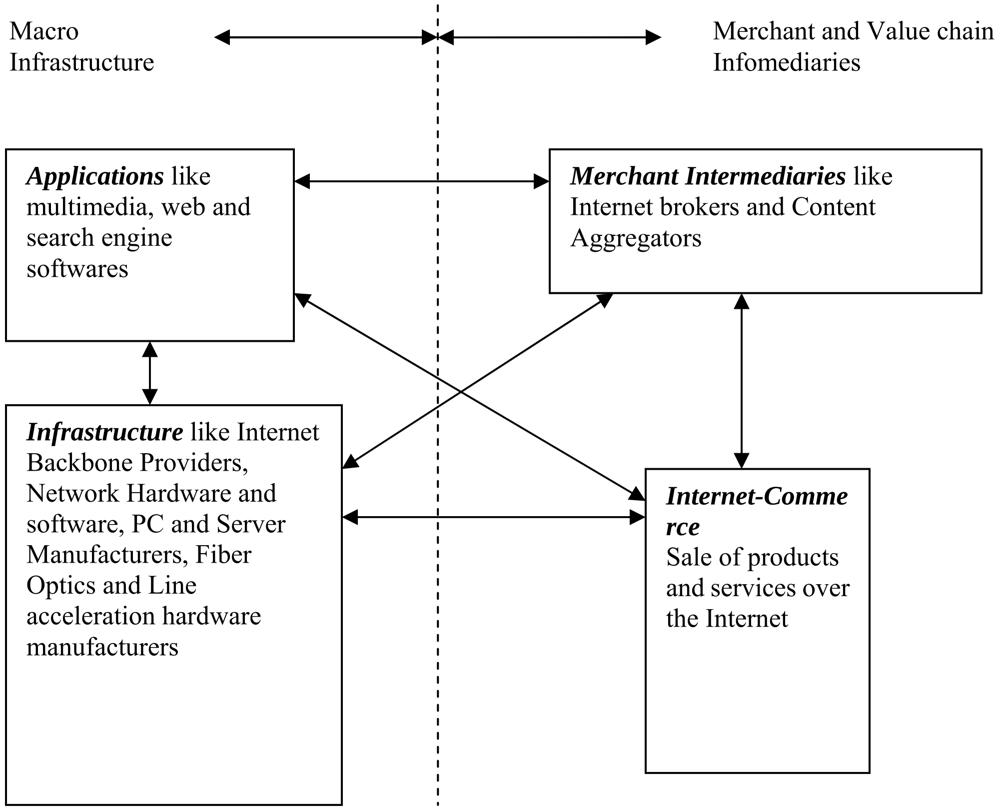

Figure 1. Segments of the Internet Economy

3. Diffusion of Innovation and its Modeling Alternatives

3.1. The Diffusion Phenomenon

The adoption of new innovations or technologies generally does not progress in a smooth linear fashion, and studies of the phenomenon have focused on explaining the reasons for this. The diffusion rate is defined as the speed at which members of a social system adopt the innovation. Innovation diffusion models make two key assumptions regarding the phenomenon. First, the existing number of adopters (of the innovation) positively drives the rate of growth. Second, the difference between the potential number of adopters at the saturation level and the number of existing adopters also influences the rate of growth. Two basic diffusion theories have been used in the literature to explain the logic behind these models. First, the *diffusion of innovation theory* [Rogers 1983], which studies diffusion as a process by which an innovation is communicated internally over time among the members of a social or market system. That is, the diffusion is caused by internal imitation or contagion. The distribution of adopters is expected to be bell-shaped, making the cumulative number of adopters over time an S-shaped curve. This cumulative number is expected to grow slowly at first due to uncertainty about the innovation in the early phase, resulting in a relatively flat curve in the beginning of the process. If the innovation succeeds, positive feedback fuels the innovation process. The adoption rate accelerates rapidly causing a steep curve which levels off later due to saturation. The second, *utility theory of social networks* [Valente 1996], suggests that potential returns from adopting a network depends on the number of existing users. This supports the idea of network externalities and this dependence is especially strong for computer and telephone networks where the greater the number of users, the greater the value of the network. Once a critical mass of membership is reached, it motivates further adoption of the innovation. If we view e-markets as networks of buyers and sellers, the existing number of suppliers and users will affect the adoption rate. The take-off point of the S-curve represents the critical mass.

Both the theoretical perspectives mentioned above essentially attribute the diffusion rates to internal causes or influences, modeled as a function of the number of existing users and the number of potential users still in the system. There is no place in either theory for external factors that may cause the diffusion rate to have sudden spikes upward or downward. [Rai et al. 1998], in their study of the diffusion of the Internet as a whole, noted that there was a sudden upward spike in the number of Internet hosts in 1987 when NSFNET switched from 640 KB to the faster T1 lines, signaling a significant investment by the U.S. government in Internet infrastructure early in the life cycle. This was clearly an external influence on the adoption behavior. Other kinds of external influences can be brought about by government intervention in allowing universal access to any new technology (for example, governments in some countries can foster or hinder access and types of access to e-commerce). Technological innovations can support and motivate an increased diversity and intensity in e-commerce applications on the Internet. These applications may require new levels of inter-operability, standards of communication protocols, and bandwidths. Any breakthrough innovation in these areas and in the areas of transactional security, electronic money, or signatures can make innovations like e-commerce more desirable to adopters and thus can instigate an external perturbation to adoption behavior of the original innovation. The Government can set tax and tariff policies to provide incentives or disincentives to accelerate or stifle adoption of a new innovation – both actions are external influences. Detailed discussions of external influences are found in the literature – technological innovations, standards, and government [Rai et al. 1998]; government and government funded agencies such as NSF, USAID, and standards [Kahn 1994], and international and socio-cultural factors [Dutta & Roy 2004]. Although the existence of external influences are known and discussed as a practical characteristic of diffusion of an innovation, very little research has been done to model the impact of such external influences, and to forecast the adoption rates of innovations that may be affected by them.

3.2. Diffusion Models

Two mathematical specifications of S-curve models – Gompertz and logistic – were used most widely in studies of diffusion growth rates [Gurbaxani 1990]. Each model allows for a diffusion growth rate that changes over time and with an eventual slowing down to a finite or bounded saturation level.

[Rai, et al. 1998] also uses the exponential function in addition to the Gompertz and logistic functions. Exponential curves assume a constant ratio of growth rate that generally characterizes the early stage of an innovation, and do not force a reduction in growth rates like the S-curves. Indeed, their study of Internet growth data (of 13 years, from January 1981 - January 1994) reported that an exponential model, with unlimited growth potential (no saturation level), outperformed both the Gompertz and the logistic models in characterizing Internet growth. However, as the authors of the study admit, this could only be true in the early stages of the diffusion process, and the model would eventually overestimate the growth rates.

While these methods have been appropriate for modeling diffusion that follows some variant of the S-curve, they fail to capture external effects on the adoption growth rates. The functional forms of the traditional models are discussed in Appendix A.

3.3. The Neural Network Model

The ability of the NN to model complex patterns should make it ideal in dealing with disturbances in diffusion data due to external effects. The basic formulation of the NN models is discussed in appendix B. Motivated by its robust capability in capturing complex phenomenon we examine two NN based approaches that can be considered novel for the purpose of modeling diffusion of innovation: a pure NN model, as discussed above, and a hybrid model that combines the NN model with the best of the traditional ones. To find the best of the traditional models we first examine all three – Gompertz, logistic, and exponential – individually. Our motivation behind examining the hybrid model is to see whether it is beneficial to combine the logic of both modeling approaches.

3.4. The Hybrid Model

It is usually beneficial to combine forecasts from different forecasts [Russell et al. 1987]. Putting equal weights on methods to combine is not recommended. Statistician's recommendation is to weight the forecasts from different methods by the inverse of individual methods MSE [Armstrong 2001 p.423]. The method essentially minimizes the expected forecast error variances. It works best when the biases (direction of forecast – forecasting high or low) of the methods are in different direction. We, therefore, combined the two approaches (or methods) as follows:

$$NF_{M,t\epsilon T} = \sum_{m\epsilon M} (W_{m,t} * F_{m,t\epsilon T}) \quad (1)$$

Where,

C(t)	is the calibration time horizon
T(t)	is the forecast time horizon in future
m	is the forecasting method
M	is the total number of forecasting methods being combined
$F_{m,t\epsilon T}$	is the forecast from method m for tЄT
$NF_{M,t\epsilon T}$	is the new combined forecast from M number of methods for tЄT
$W_{m,t\epsilon T}$	is the weight on the forecast from method m at tЄT or W_m (since weights were not varied at tЄT) and is given by:

$$W_m = (1/MSE_{m,C}) / (\sum_{m\epsilon M} (1/MSE_{m,C})) \quad (2)$$

$$MSE_{m,C} = (\sum_{t\epsilon C} (F_{m,t\epsilon C} - y_{t\epsilon C})2)/N_C$$

$y_{t \in C}$ is the actual value of dot-com counts over calibration horizon at time t in C
N_C is the number predictions in C
$F_{m,t \in C}$ is the forecast from method m for $t \in C$

Equation (2) determines the relative weights of the two basic models in creating the hybrid model on the basis of their MSE values. The lower the MSE of a model relative to the other, the greater the weight it receives. If one of the models outperforms another strongly, it may receive a weight close to 100%, discarding the weaker model.

4. Research Methodology

We used the number of hosts in the dot-com domain as a measure of the diffusion of e-commerce infrastructure growth, consistent with the literature on diffusion studies involving the Bitnet [Gurbaxani 1990] and the Internet [Press 1997; Rai et. al. 1998]. We collected the total number of live Internet sites in the dot-com domain worldwide from August 1995-July 2004 from the reports published by Netcraft on their web site www.netcraft.com [Netcraft, 2004]. Netcraft collects all the names of hosts that they can find on the Internet, and polls them with a request for the server name. In July 2001, for instance, there were 31,299,592 live sites (those that responded to electronic 'pinging'), of which 17,360,176 were dot-com sites. Since there can be multiple sites on the same machine or multiple machines used for a single popular site, Netcraft's count represents different domain names rather than the number of computers.

A two-step methodology was used to compare the predictive utility of the models. In step one, we calibrated our models – Gompertz, logistic, and exponential - on time-series data from August 1995 to July 2001. We then used the next eighteen months of data (August 2001 to January 2003) to validate the traditional diffusion models. The size of the test sample (typically about 20% of the total sample) is in line with recommendation from previous literature [Bishop 1995]. We built a multi-layered perceptron neural network model (which we will refer to as the 'pure NN' model, to distinguish it from the hybrid) to predict e-commerce growth on the same calibration sample. The best diffusion model based on its performance on both calibration and step one test samples and NN model were then selected for further analysis in step two.

In step two we combined the best step one diffusion model with the pure NN model to generate a hybrid model. We compared the forecasting performances of these three models - the best diffusion model, the pure NN model, and the hybrid model - on a larger test sample (August 2001 through July 2004), using the root mean squared error (RMSE) as a performance measure. RMSE is the preferred performance measure when different methods are compared for decision making [Armstrong & Collopy 1992; Carbone & Armstrong 1982].

5. Results and Analysis

An inspection of the dot-com host count data (Figure 2) shows the existence of external influences in e-commerce diffusion. While S-shaped curves are well-established for modeling diffusion due to internal causes, they are perhaps not sufficient to model diffusion that includes external influences. The cumulative numbers show resemblance to an S-curve, but also show enough deviation from the S-shape about half-way through, where there is a significant drop in the count.

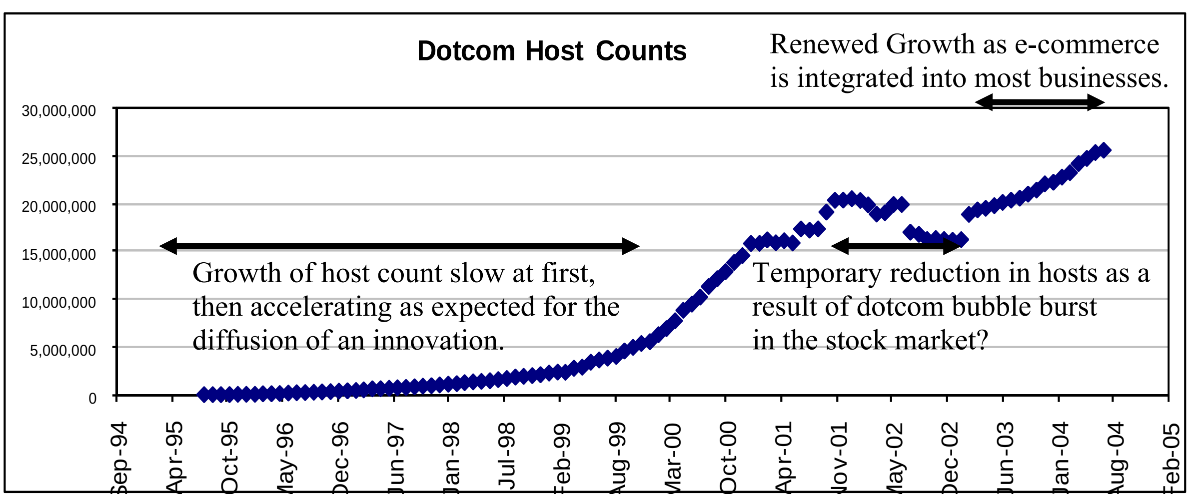

Figure 2. Growth Rate of Dotcom Hosts

Until the end of year 2000, growth seems to be exponential in nature, fitting the traditional expectations. However, towards the end of 2000 and early in the year 2001, there is actually a drop in the host count when traditional models would have forecasted continued growth, albeit at a diminished rate, to represent the upper half of the S-curve. Thus it seems that for this phenomenon, attempting the more novel NN approach might be justified.

5.1. Step 1 Results

Three traditional models, Gompertz, logistic, and exponential, were fitted to the data. Table 1 summarizes the resulting parameters. All parameters were found to be significant at a *p* value of 0.01 or better.

Table 1. Parameter Estimates of Diffusion Models and Performance on Step 1 Test Data

Model	**Parameter**	**Parameter Estimates (n=72)**	**R^2**	**Saturation Limit (SL) in # of dot-com sites**	**Inflection Point**	**RMSE on Step 1 Test Data (n=18)**
GOMPERTZ	K A B	70,861,752 0.0000022746 0.96895	0.985	70,861,752	May 2002 or 26,068,582 hosts	9,447,477
LOGISTIC	K A B	0.00000004217 0.00005910663 0.89025422613	0.991	23,696,682	September 2000 or 11,848,341 hosts	3,663,877
EXPONENTIAL	A B	2,758,485 0.2349639	0.974	No limit	None	578,176,193

$p < 0.01$ for all parameters for all 3 traditional models

The Gompertz model (R^2 equals 0.9856) predicts an approximate saturation level of 70.86 million dot-com sites. As of July 2004, the host count had reached a little over 25 million sites. The model also indicates an inflection point in May 2002. This means that the rate of dot-com site growth started to decrease sometime during the middle of 2002. The Logistic model (R^2 equals 0.9911) predicts an approximate saturation level of 23.7 million dot-com sites, with the inflection point at September 2000. The Logistic model is thus a less optimistic model than the Gompertz model in estimating the e-commerce growth. The exponential model (R^2 equals 0.9744) and has no saturation point. It is therefore, perhaps, too optimistic at this stage.

The last column in Table 1 shows the RMSE for all diffusion models on the test sample. The logistic model outperformed the others with the lowest standard error of forecasts (3,663,877). The logistic model was therefore chosen to be combined with the NN model to develop the hybrid model.

The parameter estimates for the pure NN model are shown in Figure 3. T_hu1, T_hu2 and T_hu3 are the estimated final weights from input node T to hidden units 1, 2, and 3 respectively. Hu1_ldotcom, Hu2_ldotcom, Hu2_ldotcom are the estimate weights from hidden nodes 1, 2, and 3 respectively, to the output node. The weights are sufficiently large, indicating that this model does not need any pruning. The NN model converges after 52 epochs with the objective function (sum of the absolute deviations among sample points) value of approximately 0.014. Since NN is treated as a nonparametric technique it does not provide a statistical test of the estimates with p-values. The NN model has smaller RMSE (271545) than that of the logistic model (500842) in the calibration data. The pure NN model showed a much smaller RMSE (1,983,596) than that of logistic model (3,663,877) on the step 1 test sample, indicating its superior generalization power.

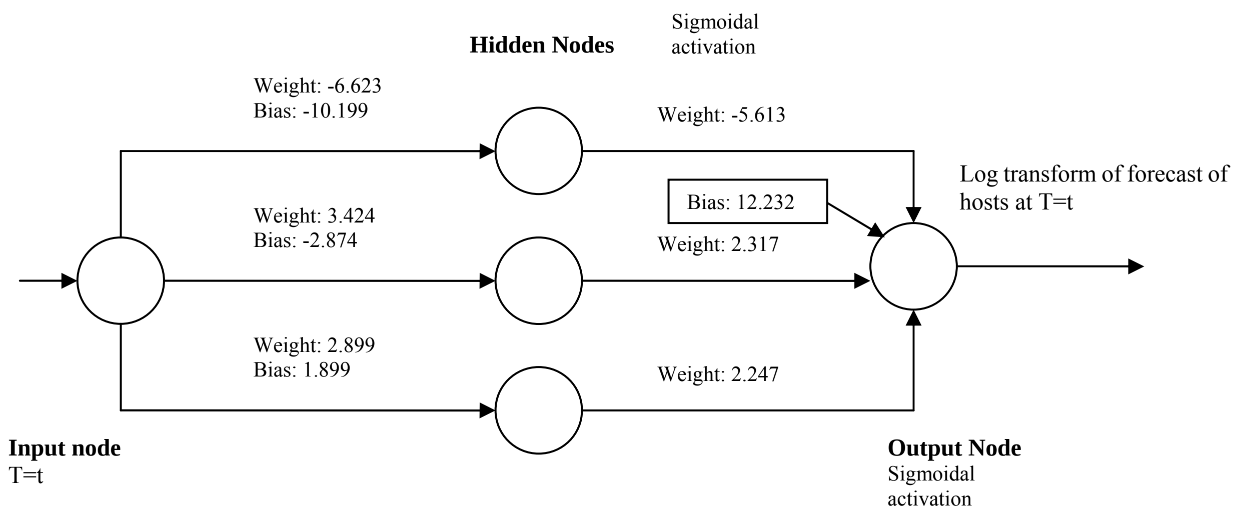

Figure 3. Parameter Estimates (weights) of the Pure NN model

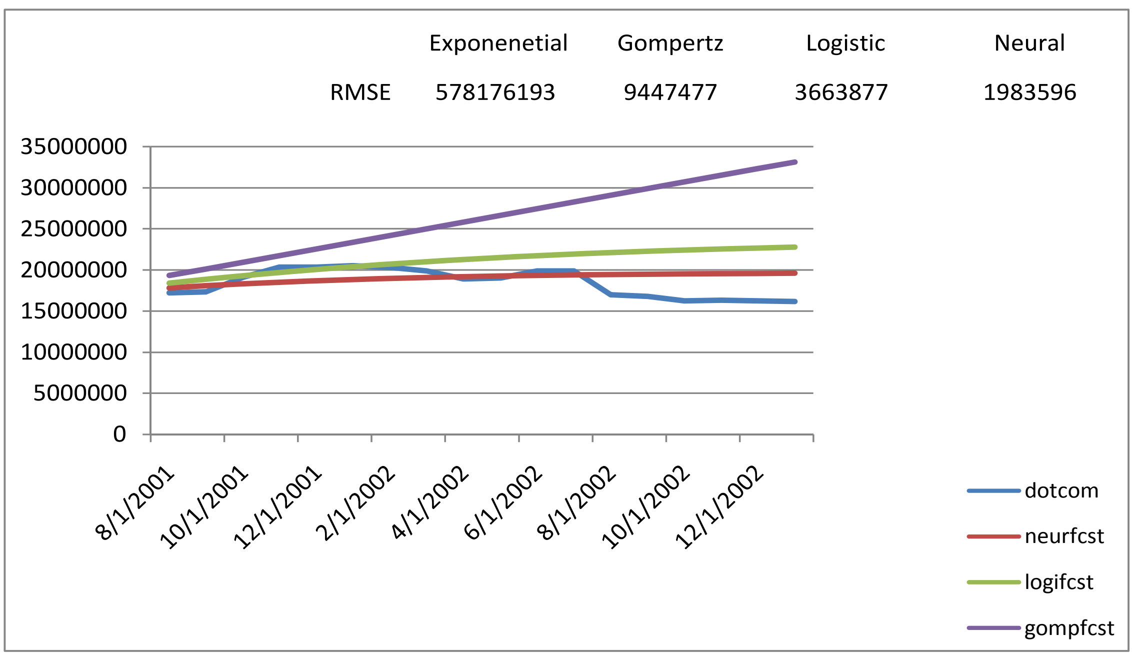

Figure 4. Forecast Performance of Models on Step 1 Test Data

Figure 4 is the graphical representation of the forecast performances of Gompertz, logistic and NN models on the test data. The graph does not contain the exponential model forecasts because of its inferior outlier performances. Among the other three models pure NN clearly outperforms other models right from the beginning of projection in future. The NN forecasts stay much closer to the actual values indicating its strength in finding the actual pattern.

5.2. Step 2 Results

The hybrid model combined the logistic and pure NN model. Based on equation (2), more weight (77%) was assigned to the pure NN model in the hybrid than to the logistic (23%), due to the lower Mean Squared Error of the former. The weight for the pure NN was computed as follows:

Weight = $[1/(271{,}545)^2]/[1/(500{,}842)^2 + 1/(271545)^2] = 0.77$.

Table 2. Forecasts and Performances on Test Data

Time Period	Actual	Neuro Forecast	Logistic Forecast	Hybrid Forecast
Aug-01	17,235,786	17,824,311	18,384,238	17,951,513
Sep-01	17,365,774	18,090,604	18,847,962	18,262,657
Oct-01	19,132,713	18,322,692	19,280,929	18,540,381
Nov-01	20,366,078	18,524,227	19,683,467	18,787,579
Dec-01	20,384,282	18,698,676	20,056,238	19,007,082
Jan-02	20,531,644	18,849,268	20,400,183	19,201,599
Feb-02	20,348,680	18,978,960	20,716,461	19,373,679
Mar-02	19,885,475	19,090,430	21,006,396	19,525,692
Apr-02	18,916,155	19,186,073	21,271,426	19,659,816
May-02	19,050,774	19,268,016	21,513,061	19,778,037
Jun-02	19,920,759	19,338,133	21,732,844	19,882,155
Jul-02	19,920,759	19,398,067	21,932,320	19,973,790
Aug-02	17,020,981	19,449,251	22,113,011	20,054,395
Sep-02	16,801,016	19,492,929	22,276,395	20,125,267
Oct-02	16,262,099	19,530,176	22,423,894	20,187,560
Nov-02	16,351,780	19,561,923	22,556,858	20,242,301
Dec-02	16,257,170	19,588,968	22,676,564	20,290,397
Jan-03	16,182,460	19,611,999	22,784,207	20,332,649
Feb-03	16,246,764	19,631,605	22,880,899	20,369,768
Mar-03	18,881,867	19,648,290	22,967,674	20,402,376
Apr-03	19,352,895	19,662,487	23,045,481	20,431,023
May-03	19,522,713	19,674,564	23,115,194	20,456,194
Jun-03	19,801,741	19,684,836	23,177,612	20,478,312
Jul-03	20,140,894	19,693,571	23,233,464	20,497,751
Aug-03	20,374,593	19,700,999	23,283,414	20,514,839
Sep-03	20,589,513	19,707,314	23,328,063	20,529,863
Oct-03	21,007,030	19,712,684	23,367,956	20,543,075
Nov-03	21,429,043	19,717,248	23,403,586	20,554,697
Dec-03	22,087,096	19,721,129	23,435,398	20,564,923
Jan-04	22,259,616	19,724,427	23,463,791	20,573,922
Feb-04	22,789,874	19,727,231	23,489,126	20,581,844
Mar-04	23,252,673	19,729,614	23,511,726	20,588,821
Apr-04	24,215,135	19,731,640	23,531,883	20,594,965
May-04	24,750,693	19,733,362	23,549,857	20,600,379
Jun-04	25,369,906	19,734,825	23,565,882	20,605,151
Jul-04	25,610,489	19,736,069	23,580,166	20,609,357
BIAS		**615,009**	**-2,221,408**	**-29,358**
RMSE		**2,551,650**	**3,284,572**	**2,428,921**
Weight used in hybrid		**0.77**	**0.23**	

The relative performances of the three models – logistic, pure NN and the hybrid model on the test data set (August 2001 to July 2004) are shown in Table 2. It is evident that pure NN forecasts are superior to logistic forecasts (RMSE: 2,551,650 and 3,284,572 respectively). However, the hybrid model was the best (RMSE: 2,428,921). Paired sample t-tests showed statistically significant differences in forecast accuracy ($p < .01$) between the pure NN model and the logistic, as well as between the hybrid model and the pure NN. The eta squared statistic also indicated a large effect size [Cohen 1988]. In summary, the results indicate that the pure NN model is superior to the logistic model (and hence the Gompertz and exponential models also) in forecasting e-commerce growth. The hybrid model performed the best of all.

The results seem to support the argument that the more flexible approach of a neural network model is better than the conventional models for forecasting innovation diffusion, especially when there is cause to believe that external factors perturb the diffusion phenomenon. However, this represents the results from one actual dataset. The validation of the current study is further strengthened by conducting a simulated experiment that compares the models across various datasets that are created to systematically examine the effects of different types of disturbances in the growth data. Specifically, one must consider four characteristics of the external effects: magnitude, direction, stage of the life cycle of diffusion, and type of S-curves. This research focused on the magnitude of the disturbance, keeping the other three invariant.

6. Simulation

6.1. Simulation Data

To further strengthen the validity of the new approach to model diffusion of e-commerce we generated five different variations of S-curve data, each with 100 observations. The first set (series 1) of 100 data points represent the baseline case of no external effects, thus presuming that only internal (contagion or imitation) factors affect the diffusion rate. The numbers can be seen as a fictitious set of the cumulative number of dot-com hosts in millions. Consistent with the theory regarding diffusion due to internal factors, the numbers form an S-curve with some random noise. The second set (series 2) of 100 numbers began as in the first, but had a 5% jump in the values starting around the inflexion point. The third, fourth, and fifth sets (series 3, 4, and 5 respectively) were similar to the second, but with increasing magnitudes of the spikes. They had spikes of 10%, 15%, and 20% respectively, in each case starting at the same time period. We used the first 80 points for calibration and the next 20 for validation.

6.2. Simulation Results and Analysis

The forecasts and RMSE of two of the models – logistic and pure NN – on the test samples for each of the five series are shown in Table 3. For the simulated data, the relative performance of the logistic regression compared with the NN was poor enough to give it a weight of almost 0. Thus, the forecasts of the hybrid model were essentially the same as those of the pure NN model. Table 3 therefore does not show the forecasts from the hybrid model separately.

Table 3. Forecasts and Performances of Best Models on Simulated Test Data

s1-s5: Simulated Test Data Points; LF = Logistic Forecast; **NF** = Neural Forecast															
T	**s1**	**LF**	**NF**	**s2**	**LF**	**NF**	**s3**	**LF**	**NF**	**s4**	**LF**	**NF**	**s5**	**LF**	**NF**
81	1040	1028	1053	1065	1053	1045	1091	1078	1072	1116	1104	1126	1141	1129	1150
82	1051	1035	1061	1076	1061	1054	1101	1086	1081	1127	1111	1134	1152	1136	1158
83	1062	1043	1070	1087	1068	1062	1112	1093	1089	1138	1118	1142	1163	1142	1166
84	1072	1050	1078	1097	1075	1071	1122	1099	1097	1148	1124	1150	1173	1148	1174
85	1082	1056	1085	1107	1081	1079	1132	1106	1105	1158	1130	1157	1183	1154	1181
86	1090	1063	1092	1116	1087	1086	1141	1111	1113	1166	1135	1164	1191	1159	1188
87	1099	1069	1099	1124	1093	1093	1150	1117	1120	1175	1141	1171	1200	1164	1195
88	1108	1074	1106	1133	1098	1100	1158	1122	1127	1183	1145	1177	1209	1169	1201
89	1117	1079	1112	1142	1103	1107	1167	1127	1133	1192	1150	1183	1218	1173	1207
90	1123	1084	1119	1148	1108	1114	1173	1131	1140	1199	1154	1189	1224	1177	1213
91	1129	1089	1124	1154	1112	1120	1179	1135	1146	1205	1158	1195	1230	1181	1219
92	1135	1093	1130	1160	1116	1126	1185	1139	1152	1211	1162	1200	1236	1185	1224
93	1141	1097	1135	1166	1120	1131	1191	1143	1157	1216	1165	1205	1242	1188	1229
94	1146	1101	1140	1171	1124	1137	1196	1146	1163	1222	1169	1210	1247	1191	1234
95	1151	1105	1145	1177	1127	1142	1202	1150	1168	1227	1172	1215	1252	1194	1238
96	1156	1108	1149	1181	1130	1147	1206	1153	1173	1231	1174	1219	1256	1196	1243
97	1159	1112	1154	1185	1134	1152	1210	1155	1178	1235	1177	1223	1260	1199	1247
98	1163	1115	1158	1188	1136	1157	1213	1158	1182	1239	1179	1227	1264	1201	1251
99	1166	1117	1162	1191	1139	1161	1216	1160	1187	1242	1182	1231	1267	1203	1255
100	1167	1120	1166	1193	1141	1166	1218	1163	1191	1243	1184	1235	1268	1205	1258
RMSE		**38**	**6**		**40**	**31**		**42**	**30**		**44**	**9**		**47**	**10**

The difference between the RMSE values for the NN model and the Logistic model is very high for each of the 5 simulated datasets, with the NN model outperforming the logistic in every case.

When applied to the sharp increase in the dot-com host count growth around the inflexion point in each of the simulated datasets, the logistic model smoothes out the spike by fitting the curve through it as best as possible, thus beginning to underestimate the total population of adopters past the spike point. The value of shape parameter B (equation A.2) does not adjust fast enough, which results in underestimation. Pure NN on the other hand, can adjust immediately through adaptive weight updates in weight matrix $\hat{W}$ (equation B.1) to reflect the continuation of the diffusion after the spike. This phenomenon is illustrated in Figure 5 using a portion of the series 5 data around the spike. The spike occurs at time period 48 (shown in square box). The pure NN model immediately adjusted after one period during calibration. However, the logistic model could not respond to the spike. Interestingly, the pure NN model also outperformed the logistic even on the baseline data that had random noise, but no spike.

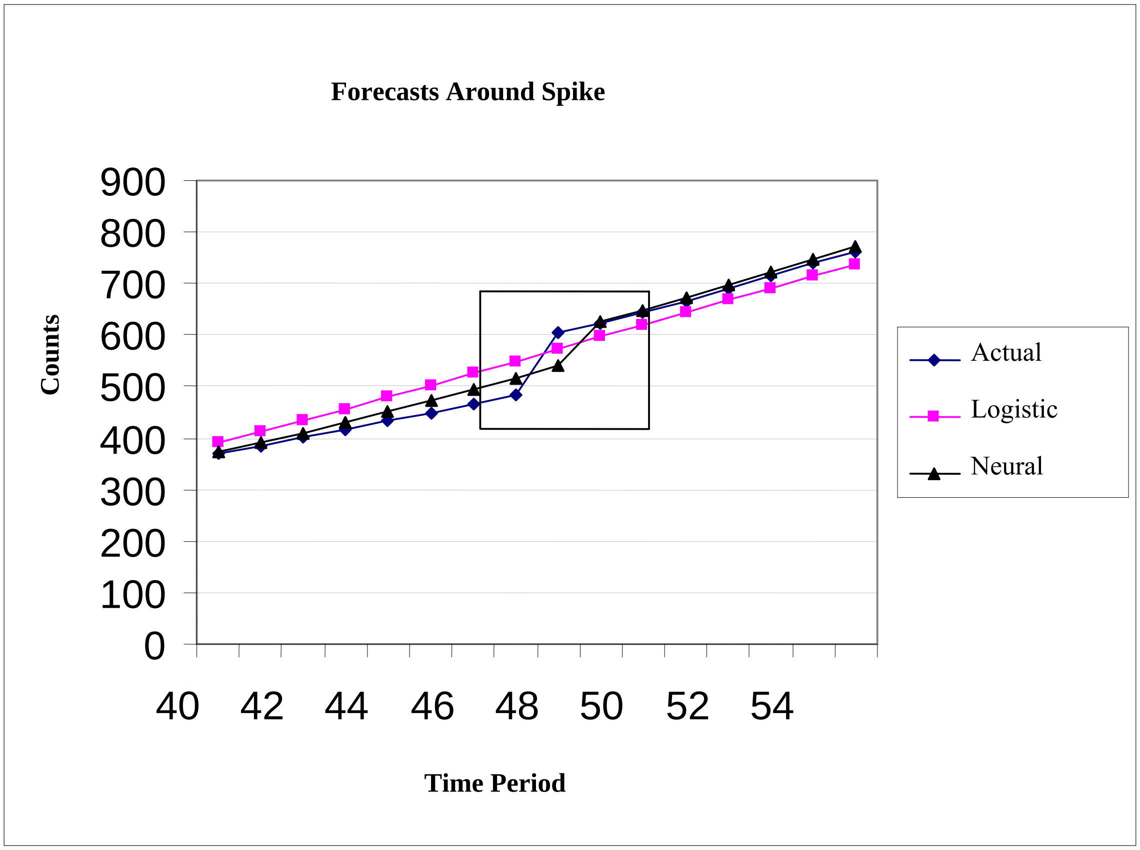

Figure 5: Logistic and Pure NN Estimates Around Spike in Series 5 Calibration Data

This ability of the pure NN model to adapt to the spike in data results in better forecasting performance on the test data, as shown in Figure 6.

The pure NN model performed extremely well compared to its traditional counterpart on simulated data. However, on the actual e-commerce data, while the pure NN model outperformed all the S-curves significantly, the hybrid model performed the best, indicating that there is still a place for the traditional models in conjunction with an adaptive NN.

We simulated "drops" (5%, 10%, 15% and 20%) in a similar way instead of "hikes." As expected, NN did better than LL on all four simulations validating NN model's superiority further in predicting e-commerce growth.

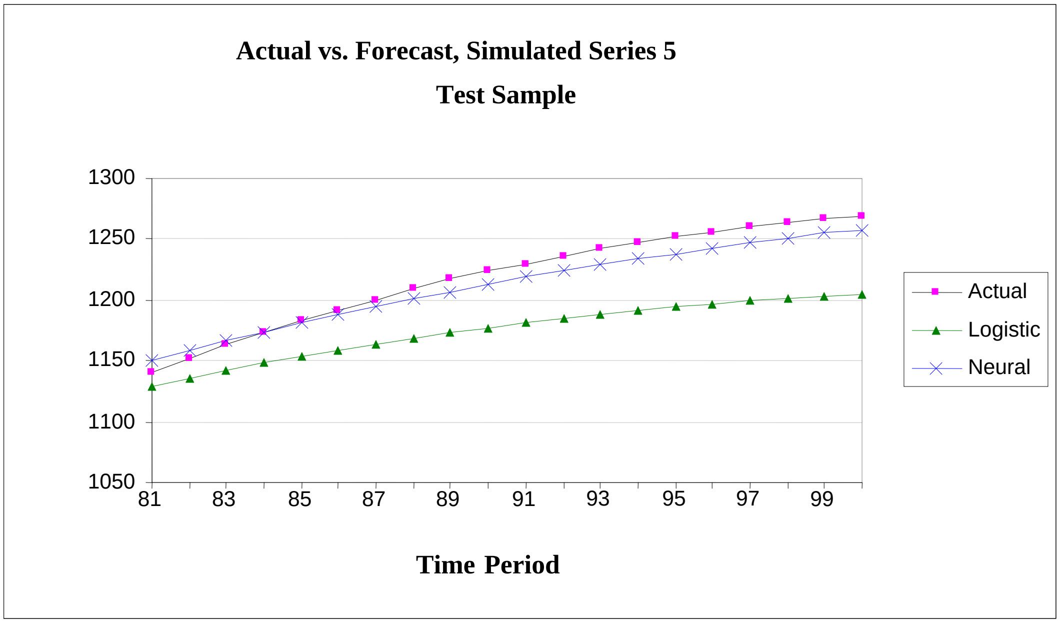

Figure 6. Logistic and Pure NN forecasts on series 5 test data

7. Conclusion

This paper makes a contribution to e-commerce research by suggesting a novel approach to model its diffusion. The internal or imitation influence is well understood in the literature. Several studies discuss the idea of imitation behavior and its effect on the diffusion of an innovation. The mathematical representations vary from models of diffusion without imitation [Coleman et al. 1966], to models with [Mansfield 1961; Bass 1969]. The imitation models are variants of S-curves. External influences are largely ignored when it comes to modeling. This is perhaps because external effects can be domain-sensitive or occur at any stage of the life cycle. Random occurrence of external effects causes problems in modeling e-commerce diffusion process with any pre-determined mathematical relation such as an S-curve. The mathematical models that fit S-curves to diffusion data thus essentially treat external perturbations as random error. These random errors reduce the accuracy of the forecasts of these models.

Our approach will be useful for most other innovation diffusion modeling studies in the future as long as there is some theoretically informed argument that external influence is suspected. This is because the modeling logic for our work here was grounded in the theories of innovation diffusion and neural networks (section 1.1). The model development was further enhanced by the simulation-theoretic approach (section 6.1). Diffusion theory motivated the search for an approach that would identify the presence of external effects in data, while the theory of neural networks provided the architectural reason why NN models are more apt to capture non-linear complexities. Additionally, the simulation experiments helped us go beyond the specific innovation of e-commerce towards more general ones. The simulations conducted in this study varied the impact of the external effects on diffusion. Neural networks performed better in each case, enabling us to argue that any innovation diffusion (not just e-commerce) that has significant external effects will be expected to benefit from this approach. Secondly, the model will hold up over time within the usual limitations of any empirically calibrated model. As [Somers 2001] points out, generalizability of NN models is assessed with the use of a test sample (a hold-out sample).

Our approach, therefore, captures internal (imitation) influences and also external influences if present. In other words, our method is suitable when the diffusion phenomenon is either a single influence or a mixed influence one. Traditional models are at best weak mixed influence models. Consequently, one would expect the forecasting performance of an NN based model for any innovation would either match or when external influence factors are present, surpass the traditional models. However, this claim needs to be tested for different innovations in the future, and will help refine the theory of innovation diffusion. When an NN model surpasses a traditional model it will signal that in all likelihood an external influence is present, which requires further investigation. Conversely, if managers and researchers dealing with e-commerce and other diffusion processes foresee that a particular innovation is especially susceptible to external influences, then they will find the results of this research extremely useful in anticipating and understanding such innovation diffusions. Overall, the pure NN method shows a lot of promise in forecasting diffusion and understanding the contributing factors.

E-commerce can be viewed as an IT innovation adopted by organizations. Economic theory suggests that organizations will adopt an innovation which cannot be protected from the competitors (by using some form of patent or other market barriers) when the innovation brings in increased profit and does not increase much risk. Additionally, as the adopters of the innovation start benefiting from the innovation, their superior performance can give reasons to others to imitate and adopt the innovation [Mansfield 1961, Rogers 1983]. However, such theoretical expectations may not necessarily come true in all actual innovation diffusions. For instance, [Mahajan et al. 1988] reported only white noise (thus no imitation influence) explaining adoption of a particular form of organizing. They concluded that the superior performance due to the innovation was not effectively visible to other organizations. The other organizations were either not convinced of the causal connection between the innovation and the superior performance of the adopter, or the innovation was too difficult to adopt. Some innovations are also subject to external perturbations which make the innovation phenomenon more complex to understand. Thus to develop a fuller understanding of why a particular innovation has spread and how it would spread in the future, each innovation has to be assessed anew to identify and study the types of influences that have caused its spread.

The presence of imitation influence indicates that e-commerce as an innovation did not suffer from lack of visibility. There were imitating adopters who expected a causal connection between e-commerce adoption and superior organizational performance. Adopting e-commerce as a way of doing business was not perceived to be challenging to organizations. This supports the notion that organizations adopted e-commerce to mimic other successful adopters. This, however, is not the complete picture. External factors also made a difference in the way e-commerce grew. Such factors include governmental investments, tax policies (e.g., no sales tax on the internet transactions); market movements that inject or take away capital infusion to organizations; interoperability and IT standards. The finding that e-commerce diffusion, influenced by significant external influences, should be modeled as a mixed-influence model phenomenon is an important and new step in e-commerce research. Similar contributions helped guide future research in other fields. In the human resource management [Somers 2001] concludes in his paper that although the common view of NN is that of predictive improvement only, his study shows that it can in fact be used to develop a new understanding of the relationship between work attitudes and job performance. For future research in e-commerce, this means setting up experiments that are designed to separate the influence categories, identify the specific factors within each, and study their relative contributions to the growth phenomenon. All such assessments will sharpen the theoretical understanding of e-commerce diffusion and planning policies that can prioritize and target external factors.

Acknowledgment
The authors thank the editors and the reviewers for their valuable suggestions.

REFERENCES

Adya, M. and F. Collopy "How Effective are Neural Networks at Forecasting and Prediction? A Review and Evaluation," *Journal of Forecasting,* Vol. 17: 143-157, 1998.

Armstrong, J. S. "Principles of Forecasting: A handbook of Researchers and Practitioners." Armstrong, J. S. (Ed), Springer 2001.

Armstrong, J. S. and F. Collopy "Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons," *International Journal of Forecasting*, Vol. 8, 69-80, 1992.

Bass, F. M., "A New Product Growth Model for Consumer Durables," *Management Science*, Vol. 15, No. 5: 215-227, 1969.

Barnden, J. A., "Simulated Reasoning, Common-sense Psychology and Artificial Intelligence," *Simulation: Evaluations and Applications*, Davis, M. and Stone, T. (eds.) Oxford: Blackwell, pp. 247-273, 1995.

Bishop, C.M., "Neural Networks for Pattern Recognition," Oxford, England: Oxford, 1995.

Carbone, R. and J.S. Armstrong, "Evaluation of Extrapolative Forecasting Methods: Results of a Survey of Academicians and Practitioners," *Journal of Forecasting*, Vol. 1:215-217, 1982.

Changsu, K. and R. D. Galliers, "Deriving a Diffusion Framework and Research Agenda for Web-based Shopping Systems," *Journal of Electronic Commerce*, Vol. 5, No. 3:199-215, 2004.

Cisco Systems, & University of Texas. (1999-2001). *Measuring the Internet Economy* from http://www.internetindicators.com/oct_1999.pdf.

Cohen , J., " Statistical Power Analysis for the Behavioral Sciences," Hillsdale, NJ: Erlbaum, 1988.

Coleman, J. S., E. K., Katz and H. Menzel, "Medical Innovation: A Diffusion Study," New-York: Bobbs-Merrill, 1966.

Dutta, A. and R. Roy, "The mechanics of Internet Growth: A Developing Country Perspective," *International Journal of Electronic Commerce*, Vol. 9, No. 2:143-165, 2004.

Elman, J.L. and D. Zipser, "Learning the Hidden Structure of Speech," *UC San Diego Institute of Cognitive Science Report 8701*, 1987.

Fourt, L.A. and J.W. Woodlock, "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing*, Vol. 25:31-38, 1960.

Golder, M. and G.J. Tellis, "Will It Ever Fly? Modeling the Takeoff of Really New Consumer Durables," *Marketing Science*, Vol. 16, No. 3:256-270, 1997.

Gurbaxani, V., "Diffusion of Computing Networks: The case of Bitnet," *Communications of the ACM,* Vol. 33, No. 12:65-75, 1990.

Hecht-Nielsen, R., "Neurocomputing," Reading, MA: Addison-Wesley, 1990.

Heeler, R. M. and T.P. Hustad, "Problems in Predicting New Product Growth for Consumer Durables," *Management Science*, Vol. 26, No. 10:1007-1020, 1980.

Kahn, R., "The Role of the Government in the Evolution of the Internet," *Communications of the ACM*, Vol. 37, No. 8:15-19, 1994.

Lippmann, R.P., "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*, 4-17, April 1987.

Mahajan, V. and E. Muller, "Innovation Diffusion and New Product Growth Models in Marketing," *Journal of Marketing*, Vol. 43, No. 4:55-68, 1979.

Mahajan, V., S. Sharma, and R. Bettis, "The Adoption of M-Form Organizational Structure: A Test of Imitation Hypothesis," *Management Science,* Vol. 34:1188-1201, 1998.

Mansfield, E., "Technical Change and the Rate of Imitation," *Econometrica*, Vol. 29, No. 4:741-766, 1961.

Netcraft 2004. Retrieved July, 2004 from http://www.netcraft.com.

Neural Network FAQ, part 3 of 7: Generalization. Retrieved July, 2004 from ftp://ftp.sas.com/pub/neural/FAQ3.html

Press, L., "Tracking the Global Diffusion on the Internet," *Communications of the ACM,* Vol. 40, No. 11:11-17, 1997.

Rai, A., T. Ravichandran, T., and S. Samaddar, "How to Anticipate the Internet's Global Diffusion," *Communications of the ACM*, Vol. 41, No. 10:97-106, 1998.

Rogers E. M., "Diffusions of Innovations," The Free Press, New York, 1983.

Roy, A. and S. Mukhopadhyay, "Interactive Generation of Higher-order Nets in Polynomial Time Using Linear Programming,' *IEEE Transactions on Neural Networks*, Vol. 8, No. 2:402- 412, March 1997.

Rumelhart, D. E., G.E. Hinton, and R.J. Williams, "Learning Internal Representations by Error Propagation," in Parallel Distributed Processing Explorations in the Microstructure of Cognition, Vol. 1, MIT Press, Cambridge, pp. 328-330, 1988.

Russell, T.D. and Everette, A. Jr. "An Empirical Evaluation of Alternative Forecasting Combinations." Management Science, October, pp. 1267-1276, 1987.

Sarle, W. S., "Stopped Training and Other Remedies for Overfitting," *Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics*, 1995.

Sejnowski, T.J. and C.R. Rosenberg, "Parallel Networks that Learn to Pronounce English Text," *Complex Systems*, Vol. 1:145-168, 1987.

Sharda, S., and R.B. Patil, "Connectionist Approach to Time Series Prediction: An Empirical Test," *Journal of Intelligent Manufacturing*, Vol. 3, No. 5:317-323, October 1992.

Somers, M.J., "Thinking Differently: Assessing Nonlinearities in the Relationship between Work Attitudes and Job Performance using Bayesian Neural Network," *Journal of Occupational and Organizational Psychology*, Vol. 74:47-61, 2001.

Tam, K.Y. and M.Y. Kiang, "Managerial Applications of Neural Networks: The Case of Bank Failure Predictions," *Management Science*, Vol. 38, No. 7:926-947, July 1992.

Valente, T.W., "Social Network Thresholds in the Diffusion of Innovations," *Social Network*, Vol. 18:69-89, 1996.

van den Bulte, C. and G.L. Lilien, "Bias and Level Diffusion Models," *Marketing Science*, Vol. 16, No. 4:338-353, 1997.

Venkatraman, N., L. Loh, and J. Koh, "The Adoption of Corporate Governance Mechanisms: A Test of Competing Diffusion Models," *Management Science,* Vol. 40:496-507, 1994.

Wang, Y.C.W., C-W. Chang, and M.S.H. Michael, "The Levels of Information Technology Adoption, Business Network, and a Strategic Position Model for Evaluating Supply Chain Integration," *Journal of Electronic Commerce Research*, Vol. 5, No. 2: 85-98, 2004.

Weiland A. and R. Leighton, "Geometric Analysis of Neural Network Capabilities," *Technical Report, Arpanet III* 385-392., 1988.

Yao, J.T., "Ecommerce Adoption of Insurance Companies in New Zealand," *Journal of Electronic Commerce Research*, Vol. 5, No. 1: 54-61, 2004.

APPENDIX A

Diffusion Model:

Traditionally, diffusion has been specified by three basic models: internal-influence, external-influence, and mixed-influence (Venkatraman et al. 1994). Mixed influence model (Bass 1969) represents both internal and external influences in the growth process:

$$dY_T/dt = (p + q\ Y_T)(K - Y_T) \qquad (1)$$

where K is the potential number of adopters of innovations, Y_T is the cumulative number of adopters at time period t, p is the coefficient of external influence, q is the coefficient of internal influence. Cumulative distribution of the mixed-influence model gives rise to a generalized logistic curve whose shape depends on the coefficients p and q.

Gompertz Model:

This is a special case of generalized logistic curve. The rate of diffusion is a function of existing adopters and the difference between the logarithms of the number of adopters at the saturation level and the existing number of adopters: $dy/dt = f\{y^*(\log y_{saturation} - \log y_{existing})\}$.

$$Y_T = KA^M \qquad (2)$$

M is equal to B^T in (1). For 0<A<1 and 0<B<1, Y_T is an increasing S-curve which reaches the upper bound or the saturation point of K (total number of adopters of the innovation) as time T approaches its theoretical limit of infinity. This curve reaches its inflection point (i.e., the point in the S-curve where the diffusion growth reaches its maximum rate and then switches from an increasing rate to a decreasing one) at Y_T=K/e (37% of its saturation level) where e is the Euler's constant.

Logistic Model:

This model has a similar structure as above except that it does not use the logarithmic form of the number of adopters. Thus, the rate of diffusion is expressed as dy/dt = f{y*(ysaturation – yexisting)}. This relation leads to the following integral form:

$$Yt = 1/(K + A^*M) \qquad (3)$$

For A>0 and 0<B<1, Y_T is an increasing S-curve which reaches the upper bound or the saturation point of 1/K as time T approaches its theoretical limit of infinity. This curve reaches its inflection point at Y_T=K/2. That is, the inflection point occurs when Y_T reaches 50% of its saturation level.

Exponential Model:

Unlike the above two, this model is characterized by a constant ratio of growth. It takes the following integral form:

$$Y_t = A + e^{BT} \qquad (4)$$

For B>0, Y_T is an ever increasing growth function that reaches infinity as T approaches its theoretical limit of infinity.

APPENDIX B

NN Model:

The equivalent nonlinear regression model form of one hidden feed forward neural network is as follows:

$$\log \hat{y}_{t+h} = \hat{\beta}_{\phi,h} + \sum_{j=1}^{n} \hat{\beta}_{j,h}\, f(I_t, \hat{w}_{h,j}) \tag{1}$$

where, I_t is the input vector of current time period value. The parameter h is the forecast horizon. $\hat{W}_{h,j}$ is the network weight vectors corresponding to forecast horizon *h* and *jth* hidden node. This research used the logistic form of transfer function *f* at each node:

$$f(I_t, \hat{w}_{h,j}) = (1+e^{-z})^{-1} \tag{2}$$

where,

$$z = \hat{w}_{h,j,\phi} + \hat{w}_{h,j}\, t.$$

These logistic activation functions (equation 2) of the multi-layer perceptron (MLP) nodes introduced nonlinearity in the model. The number of hidden nodes is *n*. There are many methods in AI literature that can be used for flexible nonlinear modeling. We used a Multi-layered Perceptron (MLPNN) trained by a back-propagation (BP) algorithm (Rumelhart *et al.,* 1988) to examine the e-Commerce growth process. We employed the stopped training method (Sarle 1995) to avoid *overfitting* or *memorization* in the calibration (training) samples.

The BP training algorithm is a supervised learning technique where the values of the independent variables along with the values of the dependent variables are fed to an MLP network input layer. The MLP network has a layer of input nodes, one or more layers of hidden nodes, and a layer of output nodes. The nodes of the first hidden layer are connected with the input layer nodes. The nodes of the output layer are connected with the last hidden layer nodes. The values assigned to connections between nodes are called "connection strengths" or "weights." The output of each node in an MLP network, sometimes called an "activation value," is a function of the inputs from the connecting nodes of the previous layer to itself and the corresponding weights. The function is called the "activation function." The outputs of the input layer nodes have the values of the input variables. The task of an MLP network is to extract the functional relationship between the input variables and the dependent variable, called "target", through proper assignments of weights. The output layer units represent the values of the target variables. An MLP uses the information available from the independent variables for each observation in the training set to compute an output value. The output value is then compared with target value to generate error signals for all units. There will be no error signal if there is no difference between the two values. Otherwise, the training involves a backward pass through the network during which error signals are sent to all units in the network. Weight changes in network connections are proportional to the error signals. In this procedure, the constant of proportionality is called the "learning rate." The larger this constant is, the larger are the changes in the weights in each step.

The main advantage of MLP over traditional diffusion models is its flexibility in creating nonlinear boundary surfaces that separate different predictive values in multidimensional input space. Nonlinearity comes through the activation functions (2) of the nodes.

Three layers of MLP neural network were used; one input layer for input variables (time), one hidden unit layer, and one output layer. Three units were used in the hidden layer (n=3 in equation 1). One output unit was used in the output layer. The input node was connected to all the hidden nodes to represent the values of the target variables. The hidden nodes were in turn connected to the output node.