

## TRACKING CHANGES IN USES AND USERS OF A WEB-BASED INFORMATION SYSTEM OVER TIME USING SERVER LOGS

By

**Ogbonnaya John Nwoha<sup>1</sup>**

Project Director  
Delta E-Commerce Connection  
Department of Agricultural Sciences  
Louisiana Tech University  
P.O. Box 10198  
Ruston, LA 71272-0045

Phone: 318-257-2303  
FAX: 318-257-4288  
E-mail: [nwoha@latech.edu](mailto:nwoha@latech.edu)

May 7, 2002

Selected Paper for  
American Agricultural Economics Association Annual Meeting  
Long Beach, California, July 28-31, 2002

© Ogbonnaya John Nwoha. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

---

<sup>1</sup> I wish to express appreciation to Sally Thompson and Michael Ward for their contributions to this paper while serving as my co-dissertation directors at the University of Illinois at Urbana-Champaign.

**Abstract**

This study analyzes the uses and users of the Strategic Soybean Information System (StratSoy). It identifies: the uses and users of StratSoy web site; patterns of use; and the determinants of use. Uses in eight content-based areas came from academia, countries, households, and businesses. After accounting for changes in consumer behavior over time, Poisson regression analysis reveals that freshness in the sense of new files, spillovers across areas, and marketing and promotion activities of site handlers increase hits to the site. Population, interest in soybean, and other motivational factors also increases the odds of hits to the web site.

**Key words:** Server logs; Poisson regression; Internet adoption and diffusion.

## **Introduction**

This paper analyzes the uses, users, and determinants of use of a web-based information system. It demonstrates that server logs can provide useful information for empirically investigating the uses and users of a web site. The Strategic Soybean Information System (StratSoy) web site located at the URL <http://www.ag.uiuc.edu/stratsoy.html> provides the server logs used in this study. The web site is dedicated to the provision of information on all aspects of the soybean industry. The site which started in the first quarter of 1995 was created to fill the perceived need of coordinating and streamlining the decision support systems across users and decision makers within the U.S. soybean organizations. The overall objective was to help the U.S. soybean organizations and the soybean industry in general become more effective, efficient, and profitable. The StratSoy site has evolved over the years to provide soybeans related information to more audiences than originally anticipated.

Web-based information systems pertaining to agriculture are becoming increasingly prevalent on the Internet. Specialty information systems on the Internet now include those dedicated to corn, pork, turkey, ratite (ostriches and emu) and so on. The costs of building and maintaining such information systems are substantial. More funding is continually requested from various sponsoring organizations by institutions and other entities for the implementation of Internet-based information systems. For instance, the first phase (March 1995 - September 1997) of the StratSoy project cost more than one half million dollars to implement. Given the cost involved in establishing and maintaining these information systems, economists are interested in understanding who is using the system, why they are using it, and what features of it they are using. This study tries to answer such questions for the StratSoy information system. Answers to

these questions could form a prelude to justifying expenditures on these sites. The answers will help policy makers make better decisions on the implementation of web sites and other information system innovations. It is speculated that other agricultural web sites or information systems share some common characteristic with StratSoy including the number of visitors to the site, the cost of establishing and maintaining the site, and the nature of information they provide.

The overall objective of this study is to analyze the adoption of StratSoy. In particular, the study identifies the uses, users, determinants of use, and patterns of StratSoy use by different user categories. It also analyzes the diffusion pattern of adoption among these user categories. The study contributes to the literature on the adoption and diffusion of a new technological innovation. It also demonstrates the application of a new type of data (server logs) to empirical investigation of the adoption phenomenon.

The findings of this research will provide the United Soybean Board (USB) which sponsors the StratSoy project with some knowledge of the returns to their investment. Such knowledge is beneficial to USB policy makers who must decide how best to allocate USB research dollars. It is hoped that this study will contribute to the existing literature on the adoption and diffusion of innovations, and particularly in the area of information technology. A study of the adoption of Internet in any walk of life is very timely since the technology is evolving at tremendous speeds and proportions. This study hopes to contribute to an understanding of the adoption and diffusion of the Internet in agriculture.

## **Uses and Users of StratSoy**

There are several possible areas to explore from the front page of StratSoy - six major and six minors.<sup>2</sup> The six major areas are labeled as: About StratSoy; Organization; Ask an Expert; Subscribe; Resources; and Database. The other six areas that one can browse or go to from the front page include; a ticker of CBOT prices, a quick search of the StratSoy server, a link to the USA weather site, Hot Picks selected by the StratSoy team, What's New area, and a contact information for the StratSoy team. For the purpose of this analysis, information available in the StratSoy site is divided into eight areas namely, the front page, research database, marketing information, production information, health and nutrition, resources, ask-an-expert, and search. The determinants of use for six of the eight areas are investigated. The production information and search areas were excluded after exploratory data analysis due to overlap with other areas.

Users of the StratSoy web site are grouped into four categories: academia, businesses, countries, and households. The number of requests (hits) from each of these user groups to the six areas of StratSoy over 48 months are analyzed. The analysis accounts for changes in consumer behavior over time, freshness in the sense of new files, spillovers across areas, and marketing and promotion activities.

## **Theoretical Framework**

The theoretical aspects of the model for understanding uses and users of the StratSoy system used in this paper derives from three major streams of theory namely, the adoption theory,

---

<sup>2</sup> The twelve areas reported here pertains to the StratSoy first page on May 28, 1998. The number of sections or areas is subject to change as new services are constantly added to the system.

the consumers' choice theory, and the diffusion theory. To the extent that information alters a priori probabilities and allows individuals to make better decisions, it is a resource that has economic value (Nicholson, 1989, p. 254 - 5). Choosing one source of information over another as presented in this section has the same underlying theory as choosing any other commodity. In particular, choosing to visit one web site over another has the same theoretical underpinning as choosing to buy one good instead of another. Users of a web site have different motivations. For instance, the first individual, a business executive, visits a web site to obtain information that enables her increase her profits. A second individual, a researcher, visits a web site to enable him to obtain information needed for his research. A third individual, an average home computer user, visits a web site to obtain nutritional information about a food item. Each of these individuals has a different motivation for visiting a web site. The theoretical model employed in this study assumes that an individual can visit a single web site over any other activity and over a specified time horizon, usually a very short time period. The individual also has the option of visiting none of the  $N$  web sites and uses his time in non-Internet related activity. The activities' choice set for the individual is

$$\{a_i\} = (a_1, a_2, \dots, a_N) \quad (1)$$

The  $a_i$  activities are mutually exclusive within a short span of time. This implies that the individual cannot be playing golf and surfing the Web at the same time. The individual would have to forego some other activity at the time he is surfing the Web. It is further assumed that the individual obtains some value from whichever activity he engages in and that there is some cost (mostly opportunity cost) associated with the activity of choice. Therefore, the value obtained

from engaging in any activity is a net value since the cost of such activity is considered. The outcome of most activities such as visiting a web site is not known until the activity is completed.

For this reason, the net value is only an expectation at the point of decision making. So, the

individual's utility maximization problem expressed in a general form is  $\frac{\max}{a_i \in \{a_j\}} NV^e(a_i)$ , (2)

where  $NV^e(a_i)$  is the expected net value obtained from choosing activity  $a_i$ . For Web

surfing we expect it to be associated with value of information. Since the visitors to a web site

come from all over the world and have different motivations, there are several stochastic

components of (2) necessitating the use of a random utility model. The random utility model

requires that the stochastic component enter the expected net value function directly. The

interpretation of these models is that although the individual may know what affects her expected

net value, the econometric investigator cannot observe all the relevant factors. These unobserved

factors may be characteristics of the individual or of the goods and may change over time (Brown

and Walker, 1989). The random utility hypothesis is appealing because it motivates the

randomness existing in an applied demand model across individuals and provides a structure for

stochastic specifications.

The individual Web surfer compares a visit to StratSoy to best other alternative activity.

Letting  $a_s$  be a visit to StratSoy, and  $a_{-s} = \frac{\arg \max}{a_i \in \{a_k\}_{k \neq s}} NV^e(a_i)$ , then

$$\frac{\max_{a_i \in \{a_j\}} NV^e(a_i)}{a_s, a_{-s}} = \frac{\max}{a_s, a_{-s}} (NV(a_{-s}), NV^e(a_s)). \quad (3)$$

Equation (3) implies that the individual compares the expected net value from visiting StratSoy with the expected value of engaging in the next best activity. He will only visit StratSoy if the expected value from such a visit exceeds his visit to any other site or engaging in any other activity. If  $D^i$  is a dummy variable for visit to StratSoy by individual  $i$ , then

$$D^i = \begin{cases} 1 & \text{iff } NV^e(a_s) \geq NV^e(a_j) \quad \forall j \neq s \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

Assume net value to be linear so that  $NV^e(a_s) = \alpha x + \varepsilon$  and  $NV^e(a_j) = \beta y + \mu$ , then

$$D^i = 1 \quad \text{iff} \quad \alpha x - \beta y + \varepsilon - \mu > 0. \quad (5)$$

Let  $R = \varepsilon - \mu$ . The researcher needs to know the distribution of  $R$  in order to obtain the estimates of  $\alpha$ 's, and  $\beta$ 's. Assume  $R$  is also normally distributed with variance  $F^2$ . The probability that an individual will visit site  $s$ , the StratSoy web site, is

$$\text{prob}(D^i = 1) = \text{prob}(\varepsilon - \mu > \beta - \alpha) = 1 - \int_{-\infty}^{\frac{\beta - \alpha}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \quad (6)$$

Individuals are not usually restricted to the number of times they can visit a site. Therefore, an individual can visit the StratSoy site as often as they wish given (5). The time period of an observation used below differs from that of activity. While activities or visits to the site can occur almost continuously, they are aggregated by months in this study. For  $N$  potential visits by the



same individual,

$$\text{prob}(\sum_{i=1}^N D^i = M; N \text{ trials}) \quad (7)$$

has a binomial distribution if the visits are independently and identically distributed (iid). This is probably a strong assumption. The probability of an individual's visit to a web site conditional on the variables available to us may be identical within say a month's time, but not likely independent. If the web site user's experience during a visit is rewarding and pleasant, that user is likely to visit the site again. In that case, subsequent visits are not independent of previous visits. There is a learning curve that occurs as an individual visits the same site over and over again such that the probability of visits is greater than zero for subsequent visits. Therefore, aggregating an individual's visit over time may violate the independence assumption.

Visits by individuals from a domain are aggregated together. While visits to the StratSoy site by individuals with similar characteristics may be independent, they are not likely identical. An aggregation of (7) over  $j$  individuals

$$\text{prob}(\sum_{j=1}^M \sum_{i=1}^N D^{ij} = MN; N \text{ trials}, M \text{ individuals}) \quad (8)$$

has a binomial distribution approximated by the Poisson distribution.

Equation (8) suggests that visits to a site aggregated over time and over individuals have the characteristics of count data. The benchmark model for count data is the Poisson regression model. Poisson regression is appropriate for dependent variables that have only nonnegative integer values: 0, 1, 2, . . . (Allison, 1999). The Poisson regression model has the underlying assumption that the dependent variable has a Poisson distribution, defined as follows. Let  $y$  be a

variable that can have only nonnegative integer values. It is assumed that the probability that  $y$  is equal to some number  $r$  is given by

$$\Pr(y = r) = \frac{\lambda^r e^{-\lambda}}{r!}, \quad r = 0, 1, 2, \dots \quad (9)$$

where  $\lambda$  is the expected value (mean) of  $y$  and  $r! = r(r-1)(r-2)\dots(1)$ . One property of the Poisson distribution is that the mean and the variance are equal, that is,

$$E(y) = \text{var}(y) = \lambda \quad (10)$$

A specification of how the parameter  $\lambda$  depends on the explanatory variables is necessary for the implementation of this model. First, write  $\lambda_{ijt}$  with the subscripts  $i$ ,  $j$ , and  $t$  to allow the parameter to vary across individuals, uses, and time ( $i = 1, \dots, n$ ,  $j = 1, \dots, 6$ , and  $t = 1, \dots, 48$ ). Since  $\lambda$  cannot be less than zero, it is standard to let  $\lambda$  be a log linear function of the  $x$  variables:

$$\log \lambda_{ijt} = \beta_{0t} + \beta_1 x_{i1t} + \beta_2 x_{i2t} + \dots + \beta_k x_{ikt} \quad (11)$$

Where  $\lambda$  is the expected value of  $y$ . It is also the mean of  $y$ . Taking logs ensures that  $\lambda$  will be greater than 0 for any values of the  $\beta$ 's or the  $x$ 's. The estimated equation is:

$$\log \lambda_{it} = \beta_{0t} + \beta_1 x_{i1t} + \beta_2 x_{i2t} + \dots + \beta_k x_{ikt} \quad (12)$$

Equation (12) is estimated 6 times (one for each use) by maximum likelihood using the GENMOD procedure of the SAS software. This equation (12) is used to estimate the betas associated with

the explanatory variables hypothesized to influence hits for academia, country, households and businesses. The beta estimates from equation (12) are interpreted as the log odds associated with the explanatory variable.

A major objective of this research is to identify the factors that influence the probability of individual  $i$  visiting the StratSoy web site. From the theory developed in this paper such factors include the characteristics of the individual as well as characteristics of the site. The net value obtained from visiting the site to an individual depends on the type of consumer, the value of the information from the site to the consumer, the characteristics of the consumer, and the characteristics of the site. The costs of visiting a web site include the opportunity cost of time, the cost of access such as the cost of computers and ISP charges, and subscription cost if the site charges a subscription fee.

One important characteristic of the StratSoy site is the availability of a large amount of information that is likely to appeal to a diverse audience. The site contains information on soybean research, soybean marketing information and data, soybean production information from around the world, and issues of health and nutrition related to the consumption of soy products. In fact, the site is a one stop shopping for all information concerning soybeans. With such a large amount of information, different areas or contents of the site are expected to appeal more to certain segment of users. Another important characteristic of the site is that fresh content is added to the site on a daily basis. Topical issues such as market conditions, health and nutrition findings, and the latest production information is posted to the site on a timely manner. In addition to the constant updates, users of the site can ask questions to different soybean subject specialists through the site.

It is hypothesized that the addition of fresh content drives more visits to the site from both repeat users and new users. The role of prior experience, subscription lists, and search engines come into play in generating hits to the site. Repeat users have come to expect the addition of new information to the site. For instance, the weekly outlook in the marketing area is updated every week and the users of that area have come to expect fresh information at a frequent interval. Other areas of StratSoy have similar content that is updated at frequent intervals. So, it is reasonable to assume that users form some expectations about the nature of updates and come back to update their knowledge from the site. Therefore, the coefficients associated with freshness of content are expected to be positive.

Another feature of StratSoy that bolsters visits by repeat users is the existence of subscription lists. Users of the site have the option of subscribing to many expert areas of the site. Questions sent to the experts and answers to those questions are copied to the mailing lists. It is expected that most users come to the site to follow up on a particular topic if it is of interest to them. While users can get every discussion on a topic in their mail box, they still come to the site to follow a thread of the topic better. The discussions are threaded in the site while they are received as single mails in the users' mail boxes. The threads are often edited by StratSoy content specialists to further add value to the information contained in this area making it more attractive for subscribers to visit the site in addition to receiving content in their mailboxes. In effect, the coefficients associated with new contents created in the expert areas are expected to be positive.

The addition of new content increases the number of new users to the site. Many new users find the site through search engines. The probability of a hit to come from a search engine increases with the amount of information available at the site at any point in time. For instance, if

Web surfers searched the entire Web for isoflavone or genistein in 1994, they had no chance of finding the StratSoy site. On the contrary, if users searched for the same words in 1996, they are more likely to find the StratSoy site because contents related to isoflavone are now available in the site. Again, the addition of new contents is expected to be positive. However, the effect on new content or freshness of information through any of these mechanisms cannot be isolated in this study.

Beyond attracting new and repeat users to the StratSoy site, the addition of new information in one area is expected to increase hits to other areas. Users visiting the site to view information in one area are likely to visit other areas of the site. This is particularly so because related information is cross linked from different areas of the site and users confidence in using the site increases as they learn how the site is organized. Therefore, the addition of fresh information in one area is expected to increase the probability of a visit to other areas of StratSoy.

Often, the addition of new information or new features to the StratSoy site is publicized through print media such as newspapers, journals, and magazines. Such publications increase users' awareness of the site and are therefore likely to increase the probability of a visit to the site by users who have access to these materials. The audiences that eventually become aware of the StratSoy site through these media depend on the nature of the media. For instance, publications made in extension bulletins are more accessible to extension agents (usually affiliated to academic institutions) and farmers. Overall, the estimated coefficients associated with publications are expected to be positive.

Users of the StratSoy web site are aggregated into four major groups based on the top level domains from which they visit the site. It is assumed that the top level domain contains

some information about the characteristics of the user including a motivation for visiting the site. The four major groups from which users visit StratSoy are academia, countries, households, and businesses. Users visiting from academic institutions are identified from the dot EDU top level domains. These users are expected to be researchers, students, or individuals affiliated to US educational institutions. Assuming that they are researchers, it is hypothesized that they are more likely to visit the StratSoy site if they are in a land grant institution. Land grant institutions are primarily research institutions and usually have a college of agriculture. Therefore they are more likely to do research on agricultural commodities such as soybeans. Furthermore, it is hypothesized that a researcher in a soybean producing state is more likely to visit the site than her counterpart in a non soybean producing state. The reasoning behind this hypothesis is that the entities that fund university research (soy related businesses, soybean organizations, and others) seem to be located in soybean producing states and are likely to fund institutions in their home state. It is further expected that institutions with higher enrolment are more likely to visit the site than those with fewer enrolments. This is the case because visits are aggregated over individuals by institutions. Over time, the cost of visiting a site is expected to drop for individuals in academia as it is dropping for users from every other user group. Even though most users from institutions do not pay for access, the costs of providing access to these users by the institutions have fallen over time making it possible for the institutions to provide more access to its members. Therefore, the number of visits from all institutions is expected to increase over time. Based on the likely motivation for use by individuals in academic institutions, they are expected to visit certain areas of the site more than other users. For instance, they are expected to be the leading users of the research database area. In comparison, visitors from the academia group are

expected to have more access to publications concerning StratSoy than other user groups.

Therefore, the coefficients associated with enrolment, land grant, and location in a soybean producing state are expected to be positive.

The second group of users of the StratSoy site comes from foreign countries, meaning countries outside of the United States of America. Since hits are aggregated over individuals by countries, it is expected that countries with larger populations will visit the StratSoy site more times than those with fewer populations, all other things being equal. Citizens of a particular country can only visit StratSoy if they have access to the World Wide Web. Therefore, citizens from a country with more Web enabling infrastructures are more likely to visit StratSoy than those with minimal or no such infrastructure. A country's interest in the soybean commodity is expected to be an important motivation for citizens of the country to visit StratSoy. The StratSoy site contained marketing and production information for the soybean commodity from all over the world and it is expected that citizens of countries that trade in soybean or produce soybeans will be interested in such information. Some countries can produce and consume soybeans domestically. If such countries are not engaged in international trade of the commodity, their motivation for visiting the site is expected to be small. Therefore, the coefficients associated with soybean trade, population, and amount of infrastructures in a country are expected to be positive.

The third group of users of the StratSoy site is households. This group comprises of users from the major Internet service providers (ISPs) such as AOL, Compuserve, and Prodigy. While users from this group may come from outside of the US, they are assumed to be primarily residents of the US. This group includes farm families and consumer households. Like academia and country users, households are expected to have a motivation for visiting the site. As

consumers of soy products, households are expected to be interested in information affecting their health and well being from soybean consumption. Since StratSoy is a web-based information technology, the households that can access this information have to know how to use the computer to a certain extent. It is expected that households' level of education will affect their ability to use the site. These households are also expected to have a certain amount of income to be able to afford a computer and the services of an ISP. The households that visit StratSoy are expected to have some interest in the soybean commodity either as consumers, producers, or marketers. Households' occupations are also expected to influence their use of the StratSoy site. For instance, farm households are expected to access production information more than non farm households.

The fourth group of users of the StratSoy site is businesses. These include businesses from both the agricultural and non agricultural industries. Agricultural businesses are expected to use the site differently than non agricultural businesses. They are more likely to be interested in an agricultural commodity than non agricultural businesses. Irrespective of the type of business, the number of employees is expected to influence the probability of visits because individuals are aggregated by businesses. Since it is assumed that business users of the site log in from their work place, only businesses that have adopted the Internet can access information from the StratSoy site. Such businesses must possess the infrastructure needed to connect to the Web. It is expected that businesses with more infrastructures are more likely to visit the site, all other things being equal. Therefore, the estimated coefficients on variables associated with the number of employees and the income of employees are expected to be positive.



## Data

The period covered by this analysis is January 1995 to December 1998. For the most part, the analysis is carried out using statistics from the web site's transfer or access server logs. Every hit on the server is recorded in the transfer log with the date and time of the request, the name or Internet Protocol (I P) address of the computer from which the request came, the actual text of the request itself, a status code, and the number of bytes transferred to the requester. An example of a raw transfer log information is reported in appendix A. The original data contain up to the second information on hits to the StratSoy server. This analysis aggregates hits on a daily, monthly, annual, or entire period depending on the level of analysis under consideration. The number of hits by each top-level domain to each area of StratSoy is extracted from the database. The original data is preserved and could be used for hourly, daily, weekly, or monthly analysis as appropriate. Table 1 reports the number of hits to StratSoy from different top level domains. An in-depth description of the raw data and other data merged to server logs is presented in Nwoha.

**Table 1. Hits to StratSoy by Top Level Domains (Jan. 1995 - Dec. 1998)**

Domain	Hits	Percent
Edu	295,222	23.50
Com	320,075	25.48
Net	258,275	20.56
Org	12,878	1.03
Mil	2,369	0.19
Gov	7,316	0.58
Others	360,220	28.67
<b>Total</b>	<b>1,256,355.00</b>	<b>100.00</b>

The web site received about 1.26 million hits between January 1995 and December 1998. About

one quarter of these hits came from 939 U.S. educational institutions, another quarter from 6149 unique dot-com domains, two fifths from 3337 dot-net domains, and about one tenth from 134 different countries. The remaining hits came from dot-org, dot-mil, dot-gov, and unreconciled domains.

About fifty-one thousand (50,664) individual files were requested through StratSoy between January 1, 1995 and December 31, 1998. Given that old files continue to receive hits years after their creation and as long as they are in the server, but less frequently, it is necessary to discount older files. A parameterized discount factor of 2 percent, 5 percent, and 10 percent is applied. This range of discount factors is consistent with literature on the depreciation of information due to advertising as published in marketing research studies (see Landes and Rosenfield, 1994 as an example). Publications and publications lagged one period are entered as binary variables in the analyses of determinants of use of the StratSoy site. The marketing activities of StratSoy personnel are measured by StratSoy expenditures. Expenditure data was obtained from StratSoy's account books. For the period of study, over half a million dollars was spent on the project (\$657,703.19). Expenditures on hosting or traveling to conferences, seminars, workshops, and farm shows accounted for about 10 percent (\$61,437.78) of StratSoy's expenditures. Monthly expenditures on travels and monthly total expenditures are merged with users monthly hits to the StratSoy site in this analysis.

For academia, usage is measured as the number of hits by individual academic institutions to the StratSoy site over time. So, the unit of analyses for academia is the institutions and time. Institutions have different characteristics that are hypothesized to influence their choice of using or not using StratSoy. Institutional characteristics used in this analysis include the following: the

type of institution derived from the institution's Carnegie classification, the size of the institution measured by enrolment, interest in the commodity measured by the intensity of soybean production in institution's state, and agricultural research in the institution measured by the land grant status of the institution.

Data used for analyzing the use of StratSoy by countries was obtained from either The World Factbook prepared by the Central Intelligence Agency (CIA) or the World Development Indicators database, World Bank, 7/1/99 prepared by the World Bank. These documents are available from their respective web sites. Data on each country's population, GDP per capita, soybean import and export, and the amounts of arable land available in the country for 1997 or 1998 were obtained. Population is an indicator of the size of country; GDP per capita is an indicator of the amount of Internet enabling infrastructures the country can afford; soybean export and import are proxies for a country's interest in the soybean commodity; and the amount of arable land is an indicator of a country's potential for soybean production. For the period of study, 134 unique countries visited StratSoy accounting for 143,585 hits to the site.

Hits to the StratSoy site from fifty-nine Internet Service Providers (ISPs) are used as a proxy for the hits from households. Several ISPs requested information from StratSoy, but analysis is restricted to ISPs with more than five hundred hits to the StratSoy site. ISP subscribers include farmers, consumers, and small businesses. These different users are lobbed together in the household category. Thirty-four of these ISPs came from the dot-net domains while the remaining 25 came from dot-com domains. Due to the difficulty in obtaining confidential information concerning users from ISPs, user characteristics could not be merged with hits to the site as done in the other user categories.

Data for businesses are obtained from the StratSoy server log files and the One Source database located at <http://www.onesource.com/>. Information on the businesses' location, annual sales, number of employees, SIC code, primary business, and primary industry are merged with the domain names and number of hits to the StratSoy site. Detailed explanations of how the data were processed could be obtained from the author.

## **Results**

Poisson regression results for the different uses of StratSoy by academia, foreign countries, households, and businesses are reported in appendix table 1 - 5. These are results from the empirical estimation of equation (12). The factors affecting the use of the StratSoy site included those relating to the site and those relating to user characteristics. For every user group, the freshness of information in the site, the marketing and promotion activities of site handlers, and the advances in technology were hypothesized to influence use. Hits to the StratSoy site by each user group were regressed on measures representing type, size, income or wealth, interest in the soybean commodity and other user characteristics in an effort to identify the determinants of use. The obtained parameter estimates are reported as percentage changes in hits associated with variables except where converting to percentage change did not make sense. In those instances, the odd ratios are reported.

## **Discussion of findings**

The study found that users are people in academia, foreign countries, households, businesses, government and non government agencies, and non profit organizations. Over 80

percent of research institutions adopted StratSoy while only about 5 percent of Associate of Arts colleges have adopted the site. More than half the entire countries of the world have visited StratSoy at least once. More than sixty thousand businesses visited the site between January 1995 and December 1998. While it was possible to merge external data with server access log data for educational institutions, countries, and businesses, the same could not be done for households. Most households access StratSoy through the services of ISPs and information provided to these ISPs are expected to remain confidential.

StratSoy was used to find soybean marketing information, production information, health and nutrition information, soybean processing information, and several other types of information about soybeans. Many of the businesses that visited StratSoy are from the software and programming industry. These users use StratSoy to acquire knowledge on new ways of delivering information to the public. Most of these companies are interested in implementing certain features of the StratSoy system such as the ask-an-expert feature, the different calculators implemented at the site, and the StratSoy tools for managing several links. So, in addition to finding soybean-related information, the StratSoy site was used as a learning resource for new ways of delivering information. Both academia and businesses used the StratSoy site for the later reason.

It was found that different factors influence different users' use of the site. The factors that determine use include those related to the site and those related to the users. Those related to the site include the freshness of information and the marketing and promotion activities of StratSoy personal. The impacts of these factors on different users differ across users.

The effect of technology advancement, captured in part by the time trend variable, was an

important determinant of use for almost all users. However, the magnitude of change in use resulting from technological cannot be clearly delineated in this study. The effect of time is confounded by other effects such as the emergence of competing web sites, and the effect of learning among users. As users become more proficient in Web surfing (learn), their pattern of use may change.

Freshness of information in terms of frequent updates was an important determinant of use for academia and foreign countries, and households. Businesses were the only group that failed to respond to fresh content. Freshness of information was an important determinant of use of health and nutrition information and marketing information areas by households. Farm households probably used marketing information more than other households. The construction of the household user category makes it difficult to dissect the different types of households. The addition of new information to one area has mixed effects for different users and uses of StratSoy. For some uses, the addition of new files to one area is associated with an increase in the use of other areas implying a cross over effect. For other uses, the addition of files in one area is associated with a decrease in use in other areas indicating a cannibalization effect. These mixed results are obtained for academia and foreign countries. The cannibalization effect was observed for households' use of the site.

Marketing and promotion activities were mostly statistically insignificant. The variable was only significant in three instances - academia's use of the front page and ask-an-expert areas, and foreign countries' use of the research database. Publications, however, were important to all users except for foreign countries and the randomly selected businesses. All the publications used in the study were domestic, hence it is not surprising that this variable had no effect on foreign

countries. Publications are expected to affect farmers' use of the site since most publications appeared in farm journals. The obtained estimates could have resulted from farmers being dwarfed by other types of households.

The study found that status as a land grant institution, location in a soybean producing state, and high enrolment are determinants of use of the StratSoy site by academic institutions. All three factors increase the probability of an institution visiting the StratSoy site. Land grant institutions are up to two times more likely to visit the StratSoy site than non land grant institutions for certain uses. Different types of educational institutions use the site differently, with the most users coming from research and doctoral institutions, and the least users coming from community colleges.

Concerning foreign countries, this study found that a country that trades in soybeans as either an exporter or an importer is more likely to use the StratSoy site than other countries. On the other hand, countries that have the potential to produce soybeans domestically, as measured by the amount of arable land, are less likely to use the StratSoy site. All things being equal, a country with high population is more likely to visit the site than other countries. The amount of technology infrastructures available to individuals in a country determines how much its citizens can use StratSoy. This study found that soybean exports, soybean imports, arable land, population and GDP per capita are important determinants of StratSoy use by foreign countries.

This study also found that the number of employees in a business and the annual sales per employee of the businesses, which are proxies for size of businesses and income of the businesses' employees, are important determinants of StratSoy use by businesses. Contrary to expectations, big businesses with many employees used the site less than businesses with fewer employees. For

agricultural businesses, those with larger number of employees used the site more than those with fewer. Similarly, sales per employee has a positive effect on use for agricultural businesses, but negative to no effect on non-agricultural business use of the site. These findings might suggest that the value of free information is higher to smaller businesses than to big businesses. However, these findings are consistent with the sample of businesses that visited the StratSoy site for the period covered in this study. Most of the businesses are from the prepackaged software industry where fewer individuals spend lots of time on the computer. Most of these businesses are start-ups with few employees and limited annual sales.

The success of StratSoy among various users demonstrates the underlying demand for quality information from a trusted source. By documenting these phenomena, this study contributes to the literature on the adoption and diffusion of a new technological innovation.

This study used statistics from the web site to study the adoption phenomena. All measures of use were obtained from the server access log files. The study therefore demonstrated the application of a new type of data to the study of adoption. This is an important contribution to economists whose analyses are often limited by the availability of data. Every web site has some access log files, and there are millions of web sites available today. The data available from these server access logs could be used to address many economic questions that have remained in the realms of theory or hypothesis.

The finding that the marketing and promotion activities of site handlers had an effect on academia but not other users has some important policy implications. It suggests that the use of a web-based information system could be increased within a user group by advertising the site to such a group. Most of the presentations that constituted the marketing and promotion activities



were made at conferences and workshops accessible to users from academia and farmers. It is most probable that marketing and promotions activities would have had an impact on households if consumer households didn't overwhelm farm households.

The results of this research indicate the importance of having fresh information on the StratSoy server. For many users, their use of an area dropped when the area was no longer updated. On the other hand, areas that are frequently updated received more hits for several user categories. It is therefore important to maintain and update a web site if users are expected to come back to the site. This implication is important in for most publicly-funded web sites. Most departments and agencies invest money in a web site, but have no specific plans for maintaining the site after a given number of years. As a matter of policy, organizations funding the establishment of web sites should make provisions for the site's maintenance. This study show that it is not prudent to build a site and leave it to fend for itself unless waning use of the site is an acceptable outcome. Publicly-funded web sites need continual support to attract and retain users of the site.

## References

Allison, P. D. *Logistic Regression Using the SAS System: Theory and Applications*. Cary, NC: SAS Institute Inc., 1999.

Brown, B. W., and M. B. Walker. "The Random Utility Hypothesis and Inference in Demand Systems." *Econometrica* 57, 4(July 1989): 815 - 829.

Landes, E. M., and A. M. Rosenfield. "The Durability of Advertising Revisited." *The Journal of Industrial Economics* XLII, 3(1994): 263-276.

Nwoha, O. J., "Empirical Analysis of StratSoy Use: Changes in Uses and Users Over Time Using Statistics from the Web Site" Doctoral thesis, University of Illinois, Urbana - Champaign, IL, May, 2001.

Nicholson, W. *Microeconomic Theory: Basic Principles and Extensions*. Fourth ed. Chicago, IL: The Dryden Press, 1989.

**Appendix Table 1. Poisson Regression Results for Academia's Use of StratSoy\***

	Front Page	Research Database	Marketing Info	Health & Nutrition	Resources	Ask-an- Expert
Parameters	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
Time	12.12					7.78
Time Squared (100)	-10.71					-9.27
Addition of New Files (100)	0.54	7.86	8.44	116.07		-37.44
New Files to Other Areas (100)	-4.57	-1.96			-0.71	0.40
Marketing & Promotion Activities (100)	-4.57					0.08
Publications Concerning StratSoy**		-0.17		0.12	-0.08	
Publications Lagged One Month**	-0.17	-0.20	-0.38	-0.07		-0.08
Enrolment (Thousands)	5.22	4.56	3.56	6.34	5.26	5.58
Interest in Soybean	2.06	2.83	2.80	0.27	1.67	1.27
Land Grant Institutions**	1.49	1.61	1.94	0.34	1.50	0.99
Master Degree Awarding** Institutions	-1.98	-1.99	-1.84	-1.81	-1.82	-1.96
Four Year Colleges**	-3.20	-3.47	-3.09	-2.83	-2.95	-2.95
Community Colleges**	-5.24	-5.74	-5.45	-4.65	-5.05	-5.00
Scale***	2.66	2.81	3.72	1.88	1.44	2.01
Number of Observations	165,663	141,524	158,775	141,511	158,773	144,957
Log Likelihood	3048	-951	-393	-3397	-4270	-2995

\* Only Statistically significant coefficients are reported (alpha =0.10). The numbers are the percentage change in hits associated with the variables except where noted.

\*\* The reported estimates for these variables are the odd ratios.

\*\*\* The scale parameter was estimated by the square root of Pearson's Chi-Squared/DOF.

**Appendix Table 2. Poisson Regression Results for Foreign Countries' Use of StratSoy\***

	Front Page	Research Database	Marketing Info	Health & Nutrition	Resources	Ask-an-Expert
Parameters	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
Time	0.02	-0.04	0.02	-0.02		
Time Squared (100)	-0.01	0.06	-0.02	0.02		
Addition of New Files (100)	0.93	20.01	11.10			
New Files to Other Areas (100)	-4.31	-0.79	-1.10	1.19		
Marketing & Promotion Activities (100)		0.12				
Publications Concerning** StratSoy						
Publications Lagged One** Month						
Population (1000 people)	0.17	0.16	0.18	0.16	0.16	0.16
GDP Per Capita (US\$1000)	7.95	7.90	7.43	9.61	7.89	8.74
Soy Imports (MMT)	19.95	19.58	18.97	8.97	16.50	12.33
Soy Exports (MMT)	4.74	2.90	3.55	2.92	3.88	2.96
Arable Land (1000 Hectares)	-1.30		-0.89	-1.24	-1.56	-0.70
SCALE***	7.58	5.16	4.78	4.11	2.84	5.15
Number of Observations	10,416	8,897	9,982	8,897	9,982	9,114
Log Likelihood	2,628	-52	744	363	-313	675

\*Only Statistically significant coefficients are reported (alpha =0.10). The numbers are the percentage change in hits associated with the variables except where noted.

\*\* The reported estimates for these variables are the odd ratios.

\*\*\* The scale parameter was estimated by the square root of Pearson's Chi-Squared/DOF.

**Appendix Table 3. Poisson Regression Results for Households' Use of StratSoy\***

	Front Page	Research Database	Marketing Info	Health & Resources Nutrition	Resources	Ask-an- Expert
Parameters	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
Time	0.02		0.02	0.03	0.01	0.01
Time Squared (100)	-18.32		-29.64	-37.96	-12.17	-17.72
Addition of New Files (100)			11.00	121.06		
New Files to Other Areas (100)	-3.24	-1.51	-1.29	-1.77	-0.63	
Marketing & Promotion Activities (100)						
Publications Concerning StratSoy**						
Publications Lagged One Month**			0.00	0.00	0.00	
Scale***	8.90	4.74	6.47	4.06	2.92	5.93
Number of Observations	48	41	46	41	46	42
Log Likelihood	8000	1420	2742	4885	3478	3452

\*Only Statistically significant coefficients are reported (alpha =0.10). The numbers are the percentage change in hits associated with the variables except where noted.

\*\* The reported estimates for these variables are the odd ratios.

\*\*\* The scale parameter was estimated by the square root of Pearson's Chi-Squared/DOF.

**Appendix Table 4. Poisson Regression Results for Randomly Selected Businesses' Use of StratSoy\***

Variables	Front Page	Research Database	Marketing Info	Health & Nutrition	Resources	Ask an Expert
Parameters	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
Time						66.53
Time Squared (100)						-0.07
Addition of New Files (100)						-0.08
New Files to Other Areas (100)						
Marketing & Promotion Activities (100)						
Publications Concerning** StratSoy						
Publications Lagged One** Month						
Number of Employees in a firm (1000)	-0.53	-0.67	-0.77	-0.37	-0.90	-0.73
Average Income of Employees in a firm (\$1000)	-19.59					-71.11
Scale***	5.34	2.55	4.18	2.40	2.08	5.28
Number of Observations	584	183	293	291	251	385
Log Likelihood	327.91	55.36	121.39	63.82	58.99	102.58

\*Only Statistically significant coefficients are reported (alpha =0.10). The numbers are the percentage change in hits associated with the variables except where noted.

\*\* The reported estimates for these variables are the odd ratios.

\*\*\* The scale parameter was estimated by the square root of Pearson's Chi-Squared/DOF.

**Appendix Table 5. Poisson Regression Results for Agriculture Related Businesses' Use of StratSoy\***

Variables	Front Page	Research Database	Marketing Info	Health & Nutrition	Resources	Ask an Expert
Parameters	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
Time		116.86				44.79
Time Squared (100)		-0.11				-0.05
Addition of New Files (100)						
New Files to Other Areas (100)	4.3e-03			4.1e-03		
Marketing & Promotion Activities (100)						
Publications Concerning StratSoy**						
Publications Lagged One Month**	0.43	0.56				
Number of Employees in a firm	2.83	2.67	2.73	3.14	2.51	3.03
Average Income of Employees in a firm	1.93	2.36	1.40	-0.04	1.31	3.01
Scale***	4.04	2.95	3.78	2.29	2.28	2.73
Number of Observations	5856	5002	5612	5002	5612	5124
Log Likelihood	-105.21	-182.59	-137.83	-210.50	-322.50	-183.74

\*Only Statistically significant coefficients are reported (alpha =0.10). The numbers are the percentage change in hits associated with the variables except where noted.

\*\* The reported estimates for these variables are the odd ratios.

\*\*\* The scale parameter was estimated by the square root of Pearson's Chi-Squared/DOF.

## Appendix A

### AN EXAMPLE OF A TRANSFER LOG ENTRY FROM STRATSOY

Here is an example of a typical transfer log entry for StratSoy<sup>3</sup>:

#### First User, First File

```
207.91.58.68 -- [10/Sep/1998:00:34:21 -0500] "GET /~stratsoy/new/ HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:22 -0500] "GET /~stratsoy/new/home.html HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:23 -0500] "GET /~stratsoy/new/images/logo.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:23 -0500] "GET /~stratsoy/new/images/q-search.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:23 -0500] "GET /~stratsoy/new/images/weather.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:24 -0500] "GET /~stratsoy/new/images/hotpicks.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:24 -0500] "GET /~stratsoy/new/images/line-v.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:24 -0500] "GET /~stratsoy/new/images/rgtarrow.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:24 -0500] "GET /~stratsoy/new/images/line.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:24 -0500] "GET /~stratsoy/new/images/1a.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:25 -0500] "GET /~stratsoy/new/images/btn-about2.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:25 -0500] "GET /~stratsoy/new/images/1c.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:25 -0500] "GET /~stratsoy/new/images/1d.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:25 -0500] "GET /~stratsoy/new/images/btn-subscribe.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:25 -0500] "GET /~stratsoy/new/images/btn-org.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:25 -0500] "GET /~stratsoy/new/images/btn-resource.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:25 -0500] "GET /~stratsoy/new/images/btn-expert.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:26 -0500] "GET /~stratsoy/new/images/btn-database.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:26 -0500] "GET /~stratsoy/new/images/1pixelclear.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:26 -0500] "GET /~stratsoy/new/images/whatsnew.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:26 -0500] "GET /~stratsoy/new/images/contactus.gif HTTP/1.0" 304 -
207.91.58.68 -- [10/Sep/1998:00:34:27 -0500] "GET /~stratsoy/new/images/1pixelclear.gif HTTP/1.0" 304 -
```

#### First User, Second File

```
207.91.58.68 -- [10/Sep/1998:00:34:34 -0500] "GET /~stratsoy/new/epv/ HTTP/1.0" 200 647
207.91.58.68 -- [10/Sep/1998:00:34:35 -0500] "GET /~stratsoy/new/epv/epvtoc.html HTTP/1.0" 200 8689
207.91.58.68 -- [10/Sep/1998:00:34:39 -0500] "GET /~stratsoy/new/epv/images/rgtarrow.gif HTTP/1.0" 200 147
```

#### Second User

```
38.27.160.189 -- [10/Sep/1998:00:41:38 -0500] "GET /stratsoy.html HTTP/1.0" 302 190
```

#### Third User

```
204.62.245.168 -- [10/Sep/1998:00:45:11 -0500] "GET /~stratsoy/soyhealth/formula.html HTTP/1.0" 200 1817
```

#### Fourth User

```
195.205.209.57 -- [10/Sep/1998:00:50:20 -0500] "GET /~stratsoy/expert/nutri.html HTTP/1.0" 200 8269
195.205.209.57 -- [10/Sep/1998:00:50:27 -0500] "GET /~stratsoy/expert/images/nutri.gif HTTP/1.0" 200 4553
```

---

<sup>3</sup>The underlined subheadings are included for clarity. They are not part of the server logs.