

Determinants of Recent Online Purchasing and the Percentage of Income Spent Online

Brendan Hannah & Kristina M. Lybecker

Department of Economics and Business, Colorado College

14 E. Cache La Poudre, Colorado Springs, Colorado, United States 80903

E-mail: hannah.brendan@gmail.com / kristina.lybecker@coloradocollege.edu

Abstract

The recent stagnation of electronic commerce highlights the need to understand contemporary online consumer behavior. E-commerce's slow growth has coincided with an explosion in the usage of Web 2.0 activities. These novel online venues have created many new channels for online retailers to reach buyers, yet these online activities have gone largely unstudied. This study incorporates current user demographics and Web 2.0 activities to dynamically model the determinants of two key measurements of recent online shopping, a recent purchase and the novel dependent variable, percentage of income spent online. Regression analysis is applied to a nationally representative 2007 survey of the U.S. online population. Determinants of a recent online purchase include, ownership of a credit card, PayPalTM account, listening to podcasts, participating in online auctions, and for the first time, female gender. In a second regression, positive determinants for the percentage of income spent online include male gender, educational attainment, online auctions, instant messaging, and online dating. Online spending increases with time online, and appears to compete with other forms of online entertainment and social networking. These results produce snapshots of contemporary online shoppers that can be used by electronic retailers to determine which product characteristics to highlight for greatest impact, and to efficiently target specific Web 2.0 activities, such as entertainment, podcast and social network websites, to develop new and robust marketing platforms.

Keywords: Online consumer behavior, Online-shopping, E-commerce, Web 2.0 marketing

1. Introduction

The number of Internet users in the U.S. has almost doubled over the last eight years to approximately 220 million in 2008 (Nielsen 2008) with over 70% purchasing online (Pew Internet 2009a). This expansion in the Internet population has coincided with an explosive increase in the online usage of Web 2.0 activities such as social networking, video sharing, and other entertainment sites. These novel online venues have created many more channels for online retailers to reach buyers. Yet yearly growth of e-commerce has fallen short of virtually all analysts' predictions for nearly a decade (Bakos 2001). The percentage of online consumers has remained steady at about 70% since late 2004 (Pew Internet, 2008b) when online buying represented 2.5% of national retail sales (U.S. Census Bureau, 2008). E-commerce has only grown about 0.6% in relation to retail sales over the last four years with revenue just over \$144 billion in 2008 (U.S. Census Bureau, 2009; 2010).

Web 2.0 channels present online retailers with novel opportunities to efficiently target potential buyers and market their goods and services. Two noteworthy trends indicate that the success of electronic retailers will depend on cultivating existing consumers: first, the proportion of consumers who never buy online remains steady, and second, over 85% of online shoppers continue to buy each year (Lohse et al., 1999; Pew Internet, 2008b). The ability of e-retailers to grow their markets is contingent on a thorough understanding of contemporary online behavior. Most published studies focus on the decision to adopt Internet shopping rather than recent purchasing behavior and utilize data from the 1990's, a period in which online buyers were predominantly older, wealthier, well-educated males (Swinyard and Smith, 2003). However, current user demographics have shifted to equal percentages of men and women buying online (Pew Internet, 2008a;b) who are younger and possess lower levels of income and education (Cumming and Kraut, 2002). This study seeks to fill this gap by analyzing more current online purchasing behavior and suggesting how online retailers can locate and appeal to potentially receptive buyers.

The contributions of this paper are twofold. First, by utilizing a more recent dataset and including Web 2.0 activities, we are able to unearth the direct links between Internet users and their online shopping decisions. Second, by examining the amount of income spent online, we can propose real world applications for these results and discuss marketing implications for reaching buyers through Web 2.0 channels. To accomplish this, we review previous literature on the determinants of online shopping and current online activities patronized by these consumers. Next, we explain how these variables provide the methodological foundation for the models

and the implementation of these models using this dataset. After presenting the findings and the limitations of the analysis, we apply the findings and suggest paths for future research.

2. Existing Literature

Online shopping has provided researchers with a rich set of questions and wealth of new information on consumer behavior and decision-making. Dating from the late 1990s, these studies provide both scholars and retailers with additional perspective on how the Internet has altered consumption and the factors that have facilitating this change. Information on the demographics, socioeconomics and consumer behavior of Internet shoppers has recently been synthesized into an Online Shopping Acceptance Model (OSAM) by Zhou et al. (2007). Online purchasing appears to be most related to convenience (Zhou et al., 2007) in addition to recreation and economic advantages (Donthu and Garcia, 1999; Korgaonkar and Wolin, 1999; Li et al. 1999; Swaminathan et al., 1999). The higher efficiency of e-commerce has reduced buyer search costs (Bakos, 1997) and produced lower prices for several online goods and services than their offline counterparts (Brown and Goolsbee, 2002; Brynjolfsson et al., 2003; Lee et al., 2003), offering the promise of products supplying good value to economic shoppers.

Online purchasing has also been facilitated by enhanced Internet accessibility as prices have steadily dropped and connections have become faster. Through this, online usage has developed into a daily part of nearly all American's lives for email correspondence and to obtain information (Nielsen 2008). In 2008, just over 90% of American home Internet consumers connected using broadband (WebSiteOptimization.com 2008). Higher connection speeds allow the Internet to be used more heavily for Web 2.0 activities, placing online gaming, instant messaging and social networking as the most time-intensive activities among broadband consumers (Nielsen/NetRatings, 2006). Between 32% and 35% of users visit blogs and social network websites such as MySpace, Facebook and others (Pew Internet 2009a). The same study showed that 28% access or download digital content, 26% participate in eBay and other online auctions, 52% watch video sites including YouTube and its derivatives, and 35% play games online (Pew Internet, 2008b). Advertising through these Web 2.0 channels allows online retailers to capitalize on the huge popularity of these novel online venues. Patronization of these activities can provide online retailers with detailed information on potential buyers' interests and allows for highly efficient marketing at lower costs but remains an unstudied aspect of online purchasing behavior.

To our knowledge, this study is the first in nearly a decade to define the factors that influence contemporary online shopping behavior. Using representative U.S. online consumer data and the current range of Internet activities, two models are developed to explain consumer determinants of online purchasing within the last year and the percentage of income spent online in the last three months, a variable that has never before been studied. There are several elements of this study that both set it apart as a unique contribution and mark it as a valuable addition to the existing literature. First, this study delineates factors that influence online spending rather than the decision to adopt or continue online shopping. The empirical work also utilizes a comprehensive dataset reflecting the newly-diverse, current demographics and online activities of online shoppers, information that can provide insight for online retailers. Since a large portion of the online population uses the Internet primarily for entertainment purposes and thus has greatly increased usage (Pew Internet, 2008b; 2009a), this study is the first to investigate the relationships between utilization of these new activities and online purchasing. The integration of demographic and socioeconomic variables with Web 2.0 activities provides the basis for developing the models presented here that describe the purchasing decisions of the current online population. These findings provide tangible evidence that can be utilized in marketing decisions, highlighting which gender-specific strategies will likely be the most successful. Finally, the study suggests that the greatest promise of online retailing, lower prices through reduced search costs, may not actually be its greatest attraction for today's consumers.

3. Methodology

3.1. The Models

Model 1. The first model of online purchasing within the last year (purchase, y_1) incorporates five categories of variables (demographic, socioeconomic, Internet usage, product perceptions and alternative activities). The dependent variable, y_1 , is a Bernoulli random variable: it has only two outcomes, purchase within the last year (1) or no purchase (0). A linear regression model does not fit a binary dependent variable such as y_1 , so a logistic regression model is used. The logistic model maps a linear combination of the predictors to the probability that the associated dependent variable equals 1 (Menard, 2002; Gelman and Hill, 2006). Y_1 is analyzed using a weighted (Note 1) logistic regression to ensure the results are nationally representative (Note 2). Table 1 provides a description and statistics of each variable.

$y_1 \sim \text{Bernoulli}(p)$

$$\text{(Equation 1)} \quad p = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{26} \beta_i x_i)}} -$$

Model 2. The second model examines the percentage of income spent online within the last three months (spending as a percentage of income, y_2) and incorporates the same five categories of variables described above. A linear regression model is applied to y_2 because its range is the real line (Gelman and Hill 2006). Equation 2 is analyzed as a weighted least squares regression.

$$\text{(Equation 2)} \quad y_2 \sim N(\beta_0 + \sum_{i=1}^{26} \beta_i x_i, \sigma^2)$$

3.2. Data Set and Sources

The regression models utilize a data set acquired through a leading, publicly traded market research company that specializes in e-commerce and online demographics (Note 3). Data are compiled from a portion of a forty-two question, closed-end survey conducted in June 2007. The questions focus on demographics, product preferences, and online behaviors, attitudes and activities. The 3,580 participants were selected by Ipsos from their U.S. online consumer panel, and the sample was balanced by demographic and behavioral characteristics derived from contemporaneous U.S. Census Bureau and research pertaining to the U.S. online population.. This dataset is particularly well-suited to this study due to the retail nature of the data collected. This data is proprietary market research and provides an accurate sample of the current online population. The survey questions were constructed to glean information on the online user's decision to purchase a good or service as well as amount spent online, information tailored to the needs of online businesses.

Several shortcomings of the data set may slightly reduce the explanatory power for each of the models. The data set presents the amount spent online and income as bracketed variables. The median amount spent online was divided by median quarterly income to determine the percentage of income spent online during the last three months. However, the use of median values from bracketed variables for age, amount spent online and income does not assure a constant standard error and can lead to lower explained variance. Ceilings on income (\$100,000+) and on the amount spent online (\$5,000+) also reduce the accuracy of the analysis, and the lack of specific information on high earning and spending outliers is expected to limit conclusions drawn here. The questionnaire also lacked key variables regarding demographics (ethnic background or location), socioeconomics (specific occupation, home ownership or number of dependents), Internet usage (user's technical aptitude or online fluency) and other intangible variables (perceptions of online buying or security in online transactions,) that would be expected to further influence online consumer behavior.

4. Results and Implications for e-Commerce

The models described here utilize representative U.S. data on 2007 online usage, patronization of Internet activities and user demographics.

4.1. Model 1

The logistic regression used for Model 1 is highly significant ($p < 0.001$) (Table 2). Table 4 provides details on the 26 variables studied, of which 12 are significant. The R-squared value of 0.237 is obtained by replicating the model for each participant and creating a predicted value (Note 4). Then a weighted, binary regression is run with the actual participant's outcome (purchase = 1 or no purchase = 0) as the dependent variable and the predicted values as the independent variable. While close to 24% of the variance is explained by the 26 independent variables studied, demographic variables alone explained only 3.3% of the variance (see Table 3).

The only significant demographic determinant of a recent online purchase is the gender of the participant ($p < 0.01$ Table 4). To our knowledge, this is the first time female gender is shown to be positively correlated with online shopping; in this case, purchasing within the last year. For an online user with low technological acceptance (Note 5), when adjusted for all other variables, females are 6.6% (37.9% for women vs. 31.3% for men are the probabilities of making on an online purchase) more likely than males to have made a recent online purchase. In contrast, for online users of high technological acceptance, females are only 0.4% more likely to have made a recent online purchase (98.9% vs. 98.5%) (Note 6). These result align with the shifting online demographics as noted by Pew Internet (2008a,b); an increasing percentage of women online and becoming more sophisticated users. Not surprisingly, this translates into a higher rate of recent online purchasing by women than men at the time of this survey. These results provide a reason for retailers to target advertising at websites heavily patronized by women.

The majority of socioeconomic variables showed positive correlations with a recent online purchase, providing verification of previous studies (Bellman et al., 1999; Donthu and Garcia, 1999; Korgaonkar and Wolin, 1999; Li et al., 1999). Even though the regression coefficient is small, household income is positively correlated with recent online buying. For a male user with low technological acceptance, an income increase from \$30,000 to

\$100,000, translates to a 15.1% (46.4% vs. 31.3%) higher probability of having purchased an item online within the last year. Male online users with low technological acceptance with a payment account such as PayPal™ are 15.0% (46.3% vs. 31.3%) more likely to have shopped within the last year. Female users with low technological acceptance have 23.4% (37.9% vs. 14.5%) higher likelihood of purchasing an item online if they possess a credit card. The results show that marketing preferentially to high earners and increasing access to credit cards or PayPal™ accounts should increase the adoption of online shopping.

The Internet usage determinants include several previously studied variables on communication and information retrieval. In contrast to the findings of previous studies (Huang, 1998; Bellman et al., 1999; Bhatnagar et al., 2000), weekly Internet usage is not significantly correlated with a recent purchase. Use of email, online research of products, use of a search engine, software downloads, and investigation of travel online is all positive, significant determinants of recent online shopping. Use of Instant Messenger for communication, a previously unstudied variable, is not significantly correlated.

Several Web 2.0 variables describing contemporary Internet activities such as shopping behavior, entertainment and social networking are also included in the regression models. Of these, participation in online auctions is the most significant determinant, followed by listening to podcasts. The remaining five variables are not significant determinants of recent online purchasing (blogs, online gaming, online dating, social networking, classified ads). These findings demonstrate that many Web 2.0 channels, such as social networking sites are not effective advertising channels to induce an online user's first purchase.

These data include measures of product perceptions in order to account for their influence on contemporary online consumer behavior. Recent online purchasing is strongly correlated with a user preference for products that are described as new and innovative while the variable for products described as good values is statistically insignificant. These findings indicate that new, dynamic product lines should highlight innovativeness in their descriptions and advertising.

Contrary to Bellman et al.'s (1999) theory of time starvation, time spent working proved insignificant in explaining recent online purchasing. In addition, given that weekly television hours is not a determinant, electronic retailers should be less motivated to advertise on television, at least to heavier television watchers. Their resources might be better spent advertising via new channels on the Internet that allow for targeted marketing, such as podcasts. Most users apparently view Internet shopping as a time saving activity, driven by convenience and the ability to find novel and hard to find items.

4.2. Model 2

Model 2 Using All Data. The second model utilizes a novel dependent variable, the percentage of income spent online in the last three months. As reported in Table 2, the regression is highly significant. Table 4 provides the coefficients and describes the significance of the 26 determinants. These variables explain 19.4% of the variance, while demographic variables account for only 2.1% (see Table 3).

The second model reveals that male gender and higher educational attainment are positively correlated with the percentage of income spent online in the last quarter. While demographic factors account for less of the variance than in the first regression, the percentage is statistically significant. Males spend 0.361% more of their total quarterly income online than do women, an increase of \$40.05 quarterly (actual; Note 7). This finding is particularly important since it is a far greater spending increase than previously reported for men who purchased \$3.15 more online annually than did women (Lohse et al., 1999). College graduation increases the percentage of total quarterly income spent online by 0.24% (model derived; Note 8) above a high school degree. These findings track previous results closely and align with the demographic hypotheses. Targeting an audience that is male and well educated should produce more revenue and allow for highly efficient marketing.

The influences of the socioeconomic determinants are somewhat different in the second regression. There is a significant, though very small, negative correlation between household income and percentage of income spent online because households with lower incomes tend to spend a larger percentage of income online, even though the actual amount spent online is less than that spent by higher income households. Possession of an online deferred payment account is a strong positive determinant of the percentage of income spent online in the last three months. In contrast, possession of an online payment account such as PayPal™ is not a significant determinant. Online purchasers using deferred payment accounts spend 1.79% more income per quarter. Although economic data are not available, online deferred purchases are relatively large (usually requiring a purchase of at least \$99) and possibly account for an increase in a user's online spending. To generate more revenue, these results suggest that online retailers should provide more methods for buyers to establish and utilize online deferred payment accounts, focusing less on PayPal™ and other similar payment systems.

Internet usage determinants encompass several correlated variables with the percentage of income spent online, positive determinants are high Internet usage and communication via Instant Messenger while email is a negative

determinant. The only significant information retrieval variable is obtaining stock quotes, which is positively correlated with the percentage of income spent online in the last three months. Within both models there are significant correlations with participation in online auctions but not with posting classified ads. The strong positive correlations observed here suggest that heavier Internet users, stock owners and those who participate in online auctions are more likely to spend a higher percentage of their income online, providing target markets for online retailers.

In the second model, all entertainment and social networking variables are determinants of percentage of income spent online. Podcasting and online dating are positive determinants while online gaming, visiting social network sites and blogging are negative determinants. These results suggest that online retailers could more profitably shift their advertising to podcasts, communication providers and online dating services.

While perception of a product as new and innovative is a positive determinant, good value is a significant negative determinant. Given these results, product lines that are novel or evolve frequently should be highlighted as such. E-retailers' marketing should align with consumers' perceptions and the Internet provides unparalleled flexibility for retailers to update websites to coincide with their target audience's perceptions and desires.

Alternative activities yield little explanatory power in either model. The lack of significant correlations between weekly television hours and percentage of income spent online again suggests electronic retailers should steer marketing projects away from television and towards more profitable advertising channels such as podcasts or other online venues that reach buyers more directly. Given that it is not currently possible to selectively bypass ads in video viewed online, advertisers have the opportunity to target marketing to this captive audience or even personalize advertising using viewers' search profiles. Although television viewing online was not a variable studied here, it should be incorporated into future research because of its rapid expansion (a sevenfold increase between 2006 and 2007) due to increased broadband adoption coupled with the virtually limitless archived and contemporaneous media available from television networks or free from third parties (Pew Internet, 2009a,b).

Model 2 Using Partitioned Data. The data set is partitioned to better understand the influence of gender, age and income, the only variables that show non-normal distributions. (Nationally representative information may still be obtained by weighting partitioned categories). The first partitioned data set is stratified by gender; then each data set is applied to the second model (Table 2). The results for men (Model 2B) show much more variance explained (31.5%) than for women (Model 2C, 8.6%). The models reveal very different determinants of online spending for men and women, but the reasons for these differences are beyond the scope of this study. The only commonalities across gender are in the impacts of income (a small, negative effect) and possession of an online deferred account, though the coefficient for men is almost six-times greater than that of women.

For males, educational attainment, deferred online payment, use of instant messaging, downloading software, online dating, and perception of new and innovative products are all significant determinants, with a deferred online payment account providing most of the predictive power (Table 4). Interestingly, significant negative determinants are use of email, online gaming and blogging. Men appear to be strongly attracted by novel products and are not value shoppers. Online gaming is a highly significant negative determinant only for men, aligning with the rest of this data suggesting that men use online shopping as a form of entertainment. Deviating from the overall data set, online male shoppers show a significant negative correlation between percent of income spent online and hours watching television.

With its lower explained variance, the regression model for women has fewer significant determinants but unearths some key gender-related differences. The determinants of women's online spending behavior support Bellman et al.'s (1999) theory of time starvation. For women only, work hours and having a high-speed connection are positive determinants. Similarly, none of the entertainment usage or social networking variables is significant. In addition to being more convenience-oriented, women who spend more online do not appear to be drawn to new or innovative products. Sophisticated Internet users, they participate in online auctions and appear to be responsible for the significance of stock quotes as a significant determinant for Model 2A.

Further partitioning of the data for men by income and age (Models 2D, 2E, 2F) yields improvements in the percentage of variability explained (Table 2). In Model 2D, for men with household incomes less than \$35,000, 66.7% of the variability is explained by this model. Positive determinants are educational attainment, deferred payment, Instant Messenger, online auctions, classified ads, and new and innovative products (Table 4). High-speed connection, product research and weekly work hours are negative determinants. A college graduate at this income level, on average, spends 1.15% (model derived) more income online quarterly than someone with just a high school degree. The most important factor, the variable with the largest coefficient, is the possession of an online deferred payment account. This result is not surprising and again points to the attention online retailers should give to such advertising venues. These results illustrate a profitable target market for retailers: men with household incomes less than \$35,000 but with a higher education seem to be novelty-oriented. Men with a

household income less than \$35,000 spend 1.17% (actual) of their income online each quarter, a higher percentage than those earning more.

Although individuals in the highest income bracket (household income of greater than \$100,000) spend more online than those in the lower income brackets, their percentage of income spent online is lower, 0.75% (actual). The partitioned data set of males with household incomes greater than \$100,000 (Model 2E) explains 29.1% of the variance in online spending (Table 2). Similar to lower earners, educational attainment, online auctions and novel products are positive determinants but not weekly work hours (Table 4). Positive Internet usage determinants are quite different for this economic group; high speed connection, weekly usage, travel research, and online dating. Not surprisingly, possession of an online deferred account is not a determinant of online spending for these highest earning men and good value is actually a negative determinant for this group. Online gaming is also a negative determinant. Given the very distinct results for high-income men, e-retailers probably require a specialized approach, using only select Web 2.0 channels to attract the business of these high earners. This group was the smallest of any examined ($n=226$), and it is possible that a model using their actual income figures might provide additional insight into the online consumer behavior of this important marketing target.

Online retailers need to rely on models that explain large amounts of variance in their target market. Model 2F, applied to men aged 45-54, accounts for 79.1% of the variance at a highly significant level despite the lack of a significant correlation with household income. In this age group, positive determinants are downloading software, instant messaging, online dating, and most significantly, having an online deferred account and products perceived as new and innovative (Table 4). The highest observed coefficients in this model are again associated with online deferred payment. Negative determinants are high-speed connection, product research, social networking sites, online gaming and weekly television hours.

While the R^2 values for Models 1 and 2 are relatively low (23.7% and 19.4%, respectively), both regression equations were statistically significant (p value < 0.001) and our determinant analysis relies only on variables that show statistical significance based on at least a p value ≤ 0.05 . Admittedly, the extent to which unknown, confounding variables affect the model's results are unknown and beyond the scope of this study. Additionally, several key variables, both demographic and socioeconomic, are missing from the data set and limit its interpretation. More complete information about survey participants would allow for a richer and more thorough analysis.

Nevertheless, the large number of survey participants allows the data to be partitioned in order to further explore purchasing behaviors. An improvement in the explained variance of Model 2 (spending) was found for men but not for women, suggesting that the inclusion of other, unstudied variables in a regression analysis may provide more explanatory power of female online shopping behavior. Partitioning the data for men enhanced the explanatory power of Model 2 for spending by men with household incomes less than \$35K and of ages 45-54. Knowledge of the causality and process between the demonstrated determinants and online spending would be necessary to translate these linkages into more effective marketing strategies.

5. Conclusions

According to the Internet Advertising Revenue Report (2009), Internet advertising grew 10.6% in 2008. In the midst of an economic downturn, and in a year in which cable television advertising was the only other category of advertising to grow, this clearly indicates the importance and potential of the online venue (Interactive Advertising Bureau 2009). Recognizing its importance, the issue then becomes how e-retailers can most effectively utilize Internet advertising, especially in the vastly growing and virtually unstudied Web 2.0 realm. The future growth and success of e-commerce may rest on the utilization of Web 2.0 platforms to advertise to specific niche users. These new conduits to potential buyers allow for the most efficient marketing to date by incorporating highly personalized and rapidly dynamic advertising based on personal interests and consumer history. This study establishes that online auctions, travel research and podcasts are strong determinants of online purchasing; they provide easily accessible advertising channels. In particular, the podcast audience is a largely untapped, growing market that may prove lucrative for online firms. Podcast consumers are at least 50% more likely than non-consumers to have made an online purchase in the past week, are avid consumers of other communication technology and are active social networkers (Webster 2008). Because possession of an online deferred payment account was the strongest determinant of spending for all groups except the highest earning men, opening a deferred payment account should be encouraged by advertising on highly correlated forums for new and innovative products, online auctions, downloadable software and podcasts.

Drawing on the results of this analysis, electronic retailers may target their efforts and consider the cost-benefit analysis of each potential advertising project. Increasing online spending in the higher income brackets, particularly by men, is crucial to the growth of electronic commerce. This demographic is well versed in

technology and has a higher than average educational attainment. It is speculated that increasing online spending in this group can be achieved through selective online advertising integrated with preferred online activities such as online auctions and possibly podcasting. Their technological sophistication allows them to block many traditional advertising methods, so advertising via podcasts and online television programs could effectively target this attractive demographic (Nesbitt 2008).

The models presented here can also provide actual revenue estimates. For example, a college graduate on average would spend 0.24% more income online than does someone with just a high school degree and increasing time online of the highest male earners by 20 hours boosts spending by an average of 0.20%, or \$55.00 per quarter for a user that makes \$100,000 per year. If online retailers could increase the percentage of income spent online by the highest income bracket, their revenues would increase significantly. Projected revenue increases can then be weighed against planned advertising outlays; the resulting cost-benefit analyses are expected to reduce marketing costs. The possibilities for targeting niche users and attraction of males to novel and innovative products have implications for website design as well.

Nearly a decade has passed since a representative population of online buyers was last studied in detail, and this research provides a more focused picture of the current American online buyer. Women are achieving parity as online shoppers as their technological savvy increases; in some cases, women utilize the Internet at a slightly higher rate than men for professional information or enrichment (Pew Internet, 2008b). However, analysis of current data provides evidence that women still appear to be time-starved purchasers. There is a small but significant increase in the amount purchased online by women as their work hours increase. Women with high-speed connections, who research products and who participate in Internet auctions spend more online. By far, though, the largest effect on spending (for both women and men) comes from having an online deferred payment account.

Although the Internet can provide novel channels to reach perfectly targeted niche users, the ability to integrate Web 2.0 activities into online marketing remains largely unexplored. The potential to customize online content and marketing for specific demographics, including by gender, by Internet needs and by online interests, is particularly important given the distinct spending patterns for men and women and the markedly different behavior of men in the highest income bracket. Marketing to new cultural or underserved ethnic groups is another route that is expected to greatly expand future e-commerce, but specific websites and advertising must be developed to accommodate cultural differences in online perceptions and types of Internet usage (Chau et al., 2002; Singh et al., 2008). With a better understanding of contemporary American online behavior, this study points to directions for future research in order to utilizing new and robust marketing platforms to reach the growing base of people who utilize the Internet for entertainment.

Acknowledgment

We thank Lauren Hannah for her assistance with the categorical statistics and Dr. Jo Ellen Hose for her editorial suggestions. We are also grateful to the Department of Economics and Business at Colorado College for financial assistance in purchasing the data.

References

- Bakos, J.Y. (1997). Reducing buyer search costs: implications for electronic marketplaces. *Management Science*, 43(12), 1676-1692.
- Bakos, Jannis. (2001). The emerging landscape for retail e-commerce. *Journal of Economic Perspectives*, 15(1), 69-80.
- Bellman, S., G.L. Lohse & E.J. Johnson. (1999). Predictors of online buying behavior. *Communications of the ACM*, 42(12), 32-38.
- Bhatnagar, A., S. Misra & H.R. Rao. (2000). On risk, convenience, and Internet shopping behavior. *Communications of the ACM*, 43(11), 98-105.
- Brown, J.R. & A. Goolsbee. (2002) Does the Internet make markets more competitive? Evidence from the life insurance industry. *Journal of Political Economy*, 110(3), 481-507.
- Brynjolfsson, E., Y. Hu & M.D. Smith. (2003). Consumer surplus in the digital economy: estimating the value of increased product variety at online bookstores. *Management Science*, 49(11), 1580-1595, 2003.
- Chau, P. Y. K., M. Cole, A. P. Massey, M. Montoya-Weiss & R.M. O'Keefe. (2002). Cultural differences in the online behavior of consumers. *Communications of the ACM*, Vol. 45(10), 138-143.

- Comiskey, D. (2006). Report: 2006 online retails sales to hit \$100 billion. [Online] Available: <http://www.e-commerce-guide.com>.
- Cummings, J.N. & R Kraut. (2002). Domesticating computers and the Internet. *Information Society*, 18(3), 221-231.
- Donthu, N. & A. Garcia. (1999). The Internet shopper. *Journal of Advertising Research*, 39(3), 52-58.
- Interactive Advertising Bureau. (2009). IAB Internet advertising revenue report. [Online] Available: http://www.iab.net/media/file/IAB_PwC_2008_full_year.pdf.
- Gelman, A & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge, UK: Cambridge University Press.
- Jupiter Research. (2007). [Ipsos online insight survey (06/07)]. Unpublished raw data.
- Korgaonkar, P.K. & L.D. Wolin. (1999). A multivariate analysis of web usage. *Journal of Advertising Research*, 39(2), 53-68.
- Lee, H.G, S.C. Lee, H.Y. Kim & R.H. Lee. (2003). Is the Internet making retail transactions more efficient?: comparisons of online and offline CD retail markets. *Electronic Commerce Research & Applications*, 2(3), 266-277.
- Li, H., C. Kuo & M.G. Russell. (2006). The Impact of Perceived Channel Utilities, Shopping Orientations, and Demographics on the Consumer's Online Buying Behavior. *Journal of Computer-Mediated Communication*, 5(2).
- Lohse, G.L., S. Bellman & E.J. Johnson. (1999). Consumer buying behavior on the Internet: findings from panel data. *Journal of Interactive Marketing*, 14(1), 15-29.
- Menard, S. (2002). *Applied Logistic Regression Analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Nie, N.H. & D.S. Hillygus.(2002). Where does Internet time come from? A reconnaissance," *IT & Society*, 1(2), 1-20.
- Nielsen. (2008). Nielsen's three screen report. [Online] Available: http://www.nielsen.com/pdf/3_Screen_Report_May08_FINAL.pdf.
- Nielsen//NetRatings. (2006). Over three-fourths of U.S. active Internet users connect via broadband at home in November, according to Nielsen//NetRatings. [Online] Available: <http://www.nielsen-netratings.com>.
- Nesbitt, A. (2008). The podcast consumer in 2008. [Online] Available: <http://www.digitalpodcast.com/podcastnews/2008/05/30/the-podcast-consumer-in-2008/>.
- Pew Internet, Inc. (2008a) Demographics of Internet users. [Online] Available: http://www.pewinternet.org/trends/User_Demo_2.15.08.htm.
- Pew Internet, Inc. (2008b) "Latest trends: usage over time. [Online] Available: <http://www.pewinternet.org/trends.asp>.
- Pew Internet, Inc. (2009a). Internet activities. [Online] Available: <http://www.pewinternet.org/Data-Tools/Download-Data/~media/infographics/Trend%20Data/January%202009%20updates/Internet%20Activities%20-%20all%20-%201%206%2009.jpg>.
- Pew Internet, Inc. (2009b). The mobile difference. [Online] Available: <http://www.pewinternet.org/Reports/2009/5-The-Mobile-Difference--Typology/2-Introduction/3-The-Typology-and-Longitudinal-Analysis.aspx?r=1>.
- Singh, N., D.W. Baack, S.K. Kundu & C. Hurtado. (2008). U.S. Hispanic consumer e-commerce preferences: expectations and attitudes toward web content. *Journal of Electronic Commerce Research*, 9(2), 162-175.
- Susskind, A.M. (2006). Electronic commerce and world wide web apprehensiveness: an examination of consumers' perceptions of the world wide web. *Journal of Computer-Mediated Communication*, 9(3).
- Swaminathan, V., E. Lepkowska-White & B.P. Rao. (1999). Browsers or buyers in cyberspace? An investigation of factors influencing electronic exchange. *Journal of Computer-Mediated Communication*, 5(2).
- Swinyard, W.R. & S.M. Smith. (2003). Why people (don't) shop online: a lifestyle study of the Internet consumer. *Psychology & Marketing*, 20(7), 567-597.
- U.S. Census Bureau. (2010). Estimated quarterly U.S. retail sales (not adjusted): total and e-commerce. [Online] Available: <http://www.census.gov/mrts/www/data/html/10Q1table4.html>.

U.S. Census Bureau. (2009). E-stats - measuring the electronic economy. [Online] Available: <http://www.census.gov/eos/www/ebusiness614.htm>

WebSiteOptimization.com. (2008). July 2008 bandwidth report. [Online] Available: <http://www.websiteoptimization.com/bw/0807/>.

Webster, T. (2008). The podcast customer revealed: The Arbitron/Edison Internet and multimedia study," Edison Media Research.

Zhou, L., L. Dai & D. Zhang. (2007). Online shopping acceptance model – A critical survey of consumer factors in online shopping. *Journal of Electronic Commerce Research*, 8(1), 41-62.

Notes

Note 1. Heteroskedasticity was not present in the results; the weighting variable is included to make the mean data and standard error for each user representative of data for the online population.

Note 2. The weight variable is based on 12 demographic variables. The idea behind the weight is to ensure that the sample is representative of the US online adult population, according to Jupiter's definition of what the demographics of that population are. Respondents are invited to take the survey based on specific demographic quotas, and then the weight is applied to the data to ensure that the distributions are precisely in line with the same demos." (Jupiter Research 2007)

Note 3. In this survey effort, Jupiter Research worked with its research partner, Ipsos Insight...Ipsos Insight is one of the largest market research companies in the US and maintains a general research panel of 400,000 households. Ipsos Insight also has access to the Ipsos US online panel, which comprises two million Internet users." (Jupiter Research 2007)

Note 4. Predicted values are determined by applying the participant's survey answers (such as age, income, high-speed connection, etc.) to the coefficients from the first regression.

Note 5. This describes an online user who is 40 years old, a college graduate, makes \$35,000 per year, works 40 hours per week, uses the Internet 5 hours per week, watches television 10 hours per week, owns a credit card, has a high-speed Internet connection and of the activities surveyed, uses only email and search engine.

Note 6. This describes an online user who is 40 years old, a college graduate, makes \$100,000 per year, works 40 hours per week, uses the Internet 20 hours per week, watches television 20 hours per week and uses the Internet for every variable surveyed.

Note 7. Actual: these values are based on descriptive statistics, in this case they are the actual differences between men and women's online spending, they are derived directly from the dataset, not from the model's results.

Note 8. Model derived: these values are obtained through the model by altering the specified variable (education) and holding all others constant.

Table 1. Description of Variables in Models and Key Statistics

DEPENDENT VARIABLE		DEFINITION OF DEPENDENT VARIABLE
y ₁	Purchase	Has purchased a good or service online within the last year (1) or not (0) n = 3580; 72% yes, 28% no
y ₂	Spending	Percentage of income spent via online purchasing in the last three months (the amount spent online in the last three months [median of bracketed range] by household income for the last three months [median of bracketed range divided by four]. n = 2539; mean [SD] = 0.682% [1.603%]
CATEGORY AND NAME OF INDEPENDENT VARIABLE		DEFINITION OF INDEPENDENT VARIABLE
Demographic		
x ₁	Age	Median age of each bracket is applied to data set: 21 (18-24), 29.5 (25-34), 39.5 (35-44), 49.5 (45-54), 55 (55+). Mean [SD] = 37.1 [13.9]
x ₂	Gender	Male (0) or female (1). 50% male, 50% female
x ₃	Educational Attainment	Highest level of education the respondent completed (1) Grade school (2) Some high school (3) Graduated high school (4) Some College (5) College graduate (6) Post-graduate degree The average respondent has completed some college. 6% some high school, 26% graduated high school, 28% some college, 28% graduated college, 12% post-graduate degree
Socioeconomic		
x ₄	Household Income	Median value from the bracketed range applied to the data set \$ 35K (<\$35K), 39.5K (35K-44K), 52K (45K-59K), 67K (60K-74K), 87.5K (75K-100K), 100K (100K+). Mean [SD] = \$59,130 [\$29,891]
x ₅	High-speed Connection	Dial-up or no connection at home (0) or high-speed (1). 63% high-speed connections, 34% dial-up connections, 2% without connection and 1% did not know their connection type
x ₆	Credit or Debit Card Ownership	Possession of credit card or debit card that can be used in online purchasing (1) or does not own any form of online payment method (0). 84% own payment card, 16% without payment card. (This variable is not used in Model 2 since an insignificant number of participants did not possess a payment card)
x ₇	Online Payment	Has online payment account such as PayPal™ (1) or not (0). 37% yes, 63% no
x ₈	Online Deferred Payment	Has online deferred payment account (1) or not (0). 5% yes, 95% no. (This variable is not used in Model 1 because it assumes user has purchased online)
Internet Usage		
x ₉	Weekly Internet Usage	Average hours spent online weekly by the user. Mean [SD] = 19.3 [21.7]
Communication		
x ₁₀	Email	User has sent or received email in last month (1) or not (0). 90% yes, 10% no
x ₁₁	Instant Messenger	User has used AOL Instant Messenger, ICQ, Yahoo or MSN Messenger, or similar instant messaging services in the last month (1) or not (0). 37% yes, 63% no
Information Retrieval		
x ₁₂	Product Research	User has researched product or services online in the last year (1) or has not (0). 52% yes, 48% no
x ₁₃	Search Engine	User searched for information using a search engine within the last month (1) or has not (0). 78% yes, 22% no
x ₁₄	Downloaded Software	User has downloaded software programs for their personal computer in the last year (1) or not (0). 25% yes, 75% no
x ₁₅	Investigated Travel	User has investigated travel arrangements (availability or pricing) online in the last month (1) or not (0). 41% yes, 59% no
x ₁₆	Stock Quotes	User has checked stock quotes online in the last month (1) or not (0). 14% yes, 86% no
Shopping Behavior		
x ₁₇	Online Auction	User has sold or bid for products in an online auction within the last month (1) or has not (0). 21% yes, 79% no
		User posted classified ads online (e.g. Craig's List, AutoTrader™, etc.) in the last month (1) or has not (0).

x ₁₈	Classified Ads	8% yes, 92% no
Entertainment		
x ₁₉	Podcast	User has listened to or downloaded a podcast within the last month (1) or not (0).
		12% yes, 88% no
x ₂₀	Online Gaming	User played games online such as action games, fantasy, flight simulators, etc. in the last month (1) or has not (0).
		27% yes, 73% no
Social Networking		
x ₂₁	Social Networks	User has visited social networking sites like MySpace™, Facebook™, etc. in the last month (1) or has not (0).
		36% yes, 64% no
x ₂₂	Online Dating	User has used an online dating service or viewed personal ads in the last month (1) or has not (0).
		15% yes, 85% no
x ₂₃	Blogs	User has read a blog in the last month (1) or has not (0).
		26% yes, 74% no
Product Perceptions		
x ₂₄	Good Value	User prefers to buy products that are a good value for the money (1) or not (0).
		78% yes, 22% no
x ₂₅	New/Innovative	User prefers to buy products that are new and innovative (1) or not (0).
		16% yes, 84% no
Alternative Activities		
x ₂₆	Weekly Work	Average hours user works weekly.
		Mean [SD] = 29.6 [27.2]
x ₂₇	Weekly Television	Average hours user watches television weekly.
		Mean [SD] = 16.9 [17.6]

Table 2. Results from Regressions

Model		Adjusted R-Squared	F Statistic (p value)	Number of Observations	Degrees of Freedom
Model 1	Purchase	0.237	1115.76 (p<0.001)	3580	3553
Model 2	Spending	0.194	23.277 (p<0.001)	2539	2512
Model 2 with Partitioned Data	Spending by Men	0.315	23.239 (p<0.001)	1213	1187
	Spending by Women	0.086	5.473 (p<0.001)	1197	1171
	Spending by Men with Household Incomes < \$35K	0.667	27.000 (p<0.001)	312	287
	Spending by Men with Household Incomes > \$100K	0.291	7.028 (p<0.001)	226	201
	Spending by Men Aged 45-54	0.791	45.250 (p<0.001)	281	256

Table 3. Results from Model #1 & #2 (Demographic Variables Only).

Model		Adjusted R-Squared	F Statistic (p value)	Number of Observations	Degrees of Freedom
Model 1	Purchase	0.033	123.242 (p<0.001)	3580	3576
Model 2	Spending	0.021	17.415 (p<0.001)	2539	2535

	Model 1 Purchase	Model 2 Spending
Constant	-1.419***	3.341***
<i>Demographic</i>		
Age	0.012***	0.001
Gender	0.217**	-1.272**
Education	0.316***	0.235*

- * 0.01 < p value ≤ 0.05
 ** 0.001 < p value ≤ 0.01
 *** p value ≤ 0.001

Table 4. Coefficients and Significance from Regressions Using Partitioned Data

	Model 1	Model 2					
		2A	2B	2C	2D	2E	2F
	Purchase	Spending	Spending by Men	Spending by Women	Spending by Men <\$35K	Spending by Men > \$100K	Spending by Men 45-54
Constant	-3.683***	0.684***	0.255	0.750***	-1.033	-1.725**	0.481
Demographic							
Age	0.006	0.002	0.007	-0.003	-0.008	0.009	◆
Gender	0.290**	-0.114*	◆	◆	◆	◆	◆
Education	0.069	0.120***	0.221***	-0.015	0.573***	0.259**	0.079
Socioeconomic							
Income	0.128***	-8E-006***	-8E-006***	-6E-006***	◆	◆	-4E-006
High-Speed Connection	0.183	0.096	0.153	0.205***	-0.429*	0.553**	-0.330*
Credit Card	1.279***	⊗	⊗	⊗	⊗	⊗	⊗
Online Payment	0.637***	-0.048	-0.155	0.040	-0.065	-0.179	0.029
Online Deferred Payment	⊗	1.797***	2.973***	0.498***	4.627***	-0.028	3.614***
Internet Usage							
Weekly Usage	0.003	0.005**	0.001	0.003	0.003	0.011*	0.001
Communication							
Email	0.588***	-0.247*	-0.516**	0.158	-0.237	0.211	-0.036
Instant Messenger	-0.133	0.235***	0.433***	-0.085	0.703***	-0.212	0.478*
Information Retrieval							
Product Research	0.471***	-0.090	-0.101	0.105	-0.542***	-0.147	-0.524**
Search Engine	0.588***	0.125	0.155	0.046	-0.391	0.093	0.119
Download	0.273*	0.112	0.211*	-0.098	0.125	0.050	0.645***
Investigate Travel	0.426***	-0.107	-0.148	0.057	-0.062	0.369*	0.054
Stock Quotes	0.215	0.162*	0.066	0.473***	-0.339	0.230	0.081
Shopping Behavior							
Online Auction	0.877***	0.125*	0.125	0.197**	0.513**	0.696***	-0.134
Classified Ads	-0.008	-0.084	0.029	-0.226*	0.798**	-0.205	0.230
Entertainment							
Podcast	0.403*	0.149*	0.086	0.089	-0.106	0.254	-0.131
Online Gaming	-0.090	-0.233***	-0.343***	-0.089	-0.192	-0.427*	-0.513*
Social Networking							
Social Networks	0.039	-0.126*	-0.129	-0.149	-0.345	-0.296	-0.743***
Online Dating	0.312	0.310***	0.336***	0.057	0.263	0.483*	0.689***
Blogs	0.318	-0.020***	-0.253**	0.100	0.260	0.101	-0.329
Product Perception							
Good Value	0.111	-0.281***	-0.320**	-0.298***	0.184	-0.656***	0.304
New/Innovative	0.391**	0.505***	0.667***	0.153	0.753***	0.439***	1.314***
Alternative Activities							
Weekly Work	-0.001	0.001	-0.001	0.003**	-0.011*	-0.006	-0.003
Weekly Television	0.005	-0.003	-0.007*	0.000	-0.006	0.002	-0.010

⊗ The variable Online Deferred Payment was not included in Model 1 because it assumes that the user has made an online purchase. The variable Credit Card was only used in Model 1 because there were an insignificant number of users in the dataset who lacked a credit card.

◆ These variables were not included when the data were partitioned by gender and/or income.

- * 0.01 < p value ≤ 0.05,
 ** 0.001 < p value ≤ 0.01,
 *** p value ≤ 0.001