Mark Lett
DOI 10.1007/s11002-013-9273-y

The relationship between consumer characteristics and willingness to pay for general online content: Implications for content providers considering subscription-based business models

Girish Punj

Abstract An increasing number of digital content providers are considering ways to charge consumers for content that was previously free. A key question for these companies is whether a change in business model from one that is advertising-based to one that is subscription-based likely to generate more revenue? Hence, the purpose of the research is to profile consumers who are more likely to pay for online content and estimate the amount they are likely to pay. Data from a nationally representative probability sample of 755 internet users are used to estimate the model. The results indicate that while the estimated *amount paid* for digital content is related to income and education, *willingness to pay* is more related to age and gender. The findings have important implications for digital content providers who are evaluating the possibility of shifting from an advertising supported content-for-free model to a subscription supported pay-for-content business model.

Keywords Digital content · Willingness to pay · Consumer characteristics · E-commerce · Online business models · Demographics

"People hate, hate, hate to subscribe to things on the internet" — Bill Gates (2005)

1 Introduction

An important feature of the web is the availability of free digital content from multiple sources. While there have been a substantial increase in the demand for online content, there

Please address all correspondence to Girish N. Punj, Department of Marketing, School of Business, University of Connecticut, 2100 Hillside Road, Storrs, CT 06269-9013. Phone: (860)-486-3835; e-mail: Girish.Punj@business.uconn.edu.

G. Punj (✉)
Department of Marketing, School of Business, University of Connecticut, 2100 Hillside Road, Storrs, CT 06269-9013, USA
e-mail: gpunj@business.uconn.edu

Published online: 03 December 2013

has been little change in consumers' willingness to pay for it. Consumers continue to regard free access to content on the internet as an inalienable right. Digital content providers, on the other hand, believe that "smart" content has the potential to change people's lives.

Companies that provide free access to their digital content mainly rely on revenues generated from advertisers, while firms that charge for access to content primarily depend on revenues generated from subscribers. An increasing number of online content providers are currently considering ways to charge consumers for content that was previously free. The shift in strategy has been prompted by declining advertising revenues from both print and digital sources. A key question for these companies is whether a change in business model from one that is advertising-based to one that is subscription-based likely to generate more revenue? The profitability of both advertising and subscription based models is driven by consumer characteristics. Companies need to understand the demographic changes that are leading to declining online advertising revenues and whether the same changes could potentially enhance the revenue streams from subscription-based models. A typical approach is to use a "freemium" business model where consumers can access some content for free but have to pay a fee for premium content. Free online content can act as a quality signal for premium content, because consumers are better able to assess its quality. Yet at the same time offering too much content for free can reduce willingness to pay for premium content. Despite the trade-off, Bourreau and Lethias (2005) analytically show that best option is to still provide some content for free, regardless of the quality of such content.

Companies that rely on advertising-based business models do so with the expectation that providing online content free increases the user base and thus enhances advertising revenue. The objective is to provide content which appeals to demographic segments valued by online advertisers. Yet, according to a 2010 survey by the Pew Internet organization, 77 % of online consumers indicated that they "hardly ever" or "never" click on an online advertisement. Similarly, in a 2009 survey conducted by the Boston Consulting Group, only half of internet users said that they would be willing to pay for online content. The survey data highlight the business challenge faced by a digital content provider considering a switch from an advertising-based model to one that relies on subscription revenues. On the one hand, consumers do not seem to be willing to pay very much for online content; while on the other, they seem inclined to ignore the advertising that enables them to receive such content for free.

A comparison of the consumer characteristics of consumers who are likely to process online advertising, with those of consumers who are likely to pay for online content, can be used to determine the likelihood of a successful transition from a content-for-free to a pay-for-content business model. Specifically, a content provider would need to know how the demographic characteristics of its users, which are of value to an online advertiser, compare with the characteristics of consumers who are likely to opt-in to a subscription-based pay model.

Hence, the purpose of the research is to profile consumers who are more likely to pay for online content and estimate the amount they are likely to pay. The research is important because the information can provide an online content provider an early prediction of the likelihood of a successful transition from an advertising-based revenue model to one that relies on subscription revenues. Information on the characteristics of consumers who are *less* willing to pay for online content can be used to implement promotional campaigns designed to break-down consumer resistance.

2 Relevant research

The monetization of online content is a formidable business challenge because the availability of free content sets the reference price for fee-based content at zero (Clemons 2009; Pauwels and Weiss 2008). Also, the transition from free to fee-based content represents a change from online content being classified as a "public good" (i.e., a non-competitive, non-exclusive resource) to a "club good" (i.e., a non-competitive, excludable resource; Buchanan 1965).

Charging for online content has been a hit-or-miss proposition, attributable to a lack of appropriate models of information value. Income is often predicted to be related to willingness to pay for online content because of its logical relationship with ability to pay. Yet, researchers have found a *negative* relationship between income and willingness to pay for certain digital products (e.g., online news content; Chyi and Yang 2009). Such a counter-intuitive result has led researchers to label such products as "inferior goods" under the theory of goods classification in microeconomics (Katz and Rosen 1991). Age and gender have been found to be related to online consumption when such content is free. An important question is how does the relationship between these characteristics change when consumers are asked to pay for such content? There is some initial evidence that many of these relationships *reverse* when consumers have to pay for content. For instance, age has been found to be negatively related to willingness to pay for online news content suggesting that younger users are more likely to pay for such content, even though they are *less* likely to be users of online news products. Likewise, males are more likely to use online news, while females are more willing to pay for it (Chyi and Yang 2009).

In sum, the findings from past research are somewhat mixed with regard to how willingness to pay for online content might be influenced by demographic factors. Yet, the degree to which consumer characteristics related to online consumption differ when it has to be paid for—as opposed to when it is free—has important implications for content providers evaluating the feasibility of shifting from a content-for-free to a pay-for-content business model.

3 Hypotheses

The present research uses a modified cost-benefit framework to understand the behavior of consumers who are more likely to pay for online content, including the amount they are likely to pay, versus seeking it out for free from alternative sources.

There are "costs" associated with locating information from either free or fee-based sources. These can be calibrated by the opportunity cost (i.e., economic value) of time spent seeking the content from a pay-for-content provider or from alternative content-for-free sources. Search costs can be expected to be higher for information from free sources. In addition there is a bundling cost associated with assimilating (i.e., bundling) information gathered from free sources. The "benefits" associated with each option can be represented by the information value such content provides them. Bundled content offers more value to consumers than corresponding unbundled content.

The opportunity cost of time is related to income. Higher-income consumers value their time more because of its opportunity cost (Stigler 1961). Thus, the amount they are likely to pay for online content is expected to be more in comparison to lower-income consumers. On the other hand, the relationship between income and willingness to pay for online content is less definite because some studies have found income to be *negatively* related to willingness to pay for online content. In addition to the effect of income, education is also likely to have an *independent* effect on the amount consumers are likely to pay for online content. Consumers with more education have a greater need for "smart" content that most pay-for-content sites provide, because they are more likely to have the expertise and cyber fluency to derive greater benefit (i.e., information value) from such content.

There are important generational and gender-based differences in the consumption of online content. For instance, age potentially affects willingness to pay for online content. Younger consumers are likely to be more willing to pay for online content as they are almost always connected to the web. More importantly, they are accustomed to paying for online products (e.g., online games, music, etc.) as they have grown up in that manner. Older consumers on the other hand have been habituated into believing that all content on the internet was intended to be free as personified by the 2005 quote from Bill Gates.

With regards to gender, women are more likely to emphasize the social aspect of information (Van Slyke, Comunale and Belanger 2002), which may increase their propensity to pay for online content. Also, they are more likely to explore a website's communication features (Jackson, Ervin, Gardiner and Schmitt 2001) and participate in virtual communities (Gefen and Ridings 2005), which may also have the same effect on willingness to pay. The above arguments lead to the two hypotheses tested in the research, namely, (a) that younger, female consumers will exhibit a greater willingness to pay for online content in comparison to older, male consumers, and (b) that the amount paid for online content will be greater for higher-income, more-educated consumers in comparison to lower-income, less-educated, consumers.

If the hypotheses are upheld, the results would indicate that the consumer characteristics associated with a greater ability to pay do not necessarily correspond to those related to a higher willingness to pay. In other words, there is a "demographic divide" between consumers based on their willingness to pay for online content and the amount they are likely to pay.

The study only examines propensity to pay for online content where they are many widely available free sources. Hence, the product category selected for this research is online news content. Other forms of digital content where there are few (if any) free sources (e.g., online music, movies, video games) are less appropriate for testing the hypotheses, because copyright protections and digital rights management (DRM) standards severely limit the use of an advertising-supported business model in these categories. Hence, our results may only be extrapolated to other general online content categories where advertising-supported and subscription-based models currently co-exist.

4 Data

Data from a national probability sample of 755 adult internet users, 18 years and older, living in the continental United States was used to test the primary hypothesis of interest in this research. The data were gathered through a telephone survey conducted by Princeton

Survey Research Associates during October 2010 on behalf of the Pew Internet & American Life Project. The non-profit sponsoring organization is an authoritative source of information on how Americans use the internet and the data provided by it is often used by federal agencies in formulating government policy, such as the recent US government policy report issued by the Federal Communications Commission (FCC) titled "Connecting America."

The survey data were collected using a dual-frame sample design. Both landline and cellular random digit dial (RDD) samples were used. The landline sample was a list-assisted random digit sample of telephone numbers selected from landline telephone exchanges in the continental USA. The cell phone sample (including those without a landline phone) was drawn from dedicated cellular exchanges based on the most recently available TPM (Terminating Point Master) data file for the continental USA. The combined sample generalizes to the American population with an internet connection, with a margin of sampling error of ±3.7 percentage points.

5 Dependent and independent variables

The first dependent variable *Pay for Online Content* was operationalized using a dichotomous scale (1=yes; 0=no) based on whether the respondent had "paid to access or download a newspaper, magazine, article or special report" either as part of a subscription or as individual file downloads. The second dependent variable, *Amount Paid for Online Content* measured the dollar amount spent by the respondent during the preceding 12 month period on accessing or subscribing to online content.

For the independent variables, *Income* was measured as the total household income from all sources before taxes in 2009 using a seven-point ordinal scale. *Education* was measured using a five-point ordinal scale. *Age* was measured using a continuous scale but then recoded into a six-point ordinal scale that used break-points in chronological age that are normally used by demographers to distinguish between generations (e.g., Gen Y, Gen X). *Gender* was recorded by the phone interviewer on a dichotomous scale. Overall, the sample distributions on the study variables closely matched the demographic profile of the American population with an internet connection, which was to be expected, due to the use of a national sample frame and probability sampling. Descriptive statistics on all study variables are reported in Table 1.

6 Preliminary analyses

A cross-tabulation between *Pay for Online Content* and the demographic variables *Income*, *Education*, *Age* and *Gender* to identify the most frequent affirmative response percentages showed that 55 % of the respondents was female, 22 % of the respondents was in the 45–54 years age category, 31 % had some college education, and 23 % was in the $75,000 to $99,999 annual household income category, as reported in Table 2. Similarly, the highest mean values for *Amount Paid for Online Content* ($ paid in the preceding 12 month period) were $171 for males, $243 for respondents in the 55–64 years age category, $209 for those in the college graduate educational category, and $200 for those in the $75,000 to $99,999 income category, as reported in Table 2.

Table 1 Descriptive statistics

	Frequency	(Percent)	Mean	(Std. Dev.)
Pay for Online Content?				
Yes	141	(19.2)		
No	593	(80.7)		
Amount Paid for Online Content				
$ in preceding 12 months			143	(297)
Income:				
Less than $20,000	74	(10.1)		
$20,000 to $29,999	68	(9.3)		
$30,000 to $49,999	139	(18.9)		
$50,000 to $74,999	98	(13.3)		
$75,000 to $99,999	102	(13.9)		
$100,000 to $149,999	64	(8.7)		
$150,000 or more	48	(6.5)		
Education:				
High school incomplete	27	(3.7)		
High school graduate	180	(24.5)		
Some college or vocational school	208	(28.3)		
College graduate	189	(25.7)		
Post graduate or advanced degree	125	(17.0)		
Age:				
18–24 years	73	(9.9)		
25–34 years	98	(13.3)		
35–44 years	111	(15.1)		
45–54 years	143	(19.5)		
55–64 years	137	(18.6)		
65+years	130	(17.7)		
Gender				
Male	330	(44.9)		
Female	405	(55.1)		

Thus, while females are more likely to *Pay for Online Content* than males (55 % versus 45 %), the mean *Amount Paid for Online Content* by males was higher in comparison to females ($171 versus $104). Similarly, while those in the 45–54 years age category level were more likely to *Pay for Online Content* than any other age category, the mean *Amount Paid for Online Content* by those in the 55–64 years age category was higher ($243 versus $81). Likewise, while respondents with some college education were more likely to *Pay for Online Content* than any other educational level category, the mean *Amount Paid for Online Content* by respondents with a college degree was higher ($209 versus $128). Thus, propensity to pay for online content does not necessarily align with the amount consumers are willing to pay, as suggested by the hypotheses.

Table 2 Cross classification of demographic characteristics with pay for online content and amount paid for online content

Demographic characteristics	Pay for online content?	Amount paid for online content in preceding 12 months
	Yes[a] (%)	Mean ($)
Income:		
Less than $20,000	6.4	46.0
$20,000 to $29,999	8.2	156.0
$30,000 to $49,999	17.3	169.3
$50,000 to $74,999	17.3	146.8
$75,000 to $99,999	22.7	199.7
$100,000 to $149,999	14.5	172.8
$150,000 or more	13.6	165.5
Education:		
High school incomplete	1.4	20.0
High school graduate	15.6	37.5
Some college or vocational school	30.5	128.4
College graduate	25.5	208.6
Post graduate or advanced degree	27.0	122.7
Age:		
18–24 years	10.9	92.1
25–34 years	14.1	145.6
35–44 years	15.6	169.7
45–54 years	21.9	81.4
55–64 years	20.3	243.0
65+years	17.2	65.6
Gender		
Male	45.4	170.8
Female	54.6	103.5

Note: Entries are column percentages

7 Model estimation

Logistic regression analysis was used to formally test the hypothesized relationships. *Pay for Online Content* was used as the dependent variable in the logistic regression equation, while the demographic factors *Income*, *Education*, *Age*, and *Gender* were entered as independent variables. The −2 log likelihood difference (1,082.38) between a null (i.e., intercept only) and the logistic regression model indicated a significant fit (χ^2=191.05; 17 *df*; p<.01) with a Cox and Snell R^2=.19 and a Nagelkerke R^2=.25, as shown in Table 3.

A review of the Odds Ratio coefficients for *Income* showed that respondents who were in the $20,000 or less income category were approximately three times more likely [(Exp (β)=3.1] to *Pay for Online Content* than those in the $150,000 or more

Table 3 Logistic regression model with pay for online content as dependent variable

	β	Wald's statistic	Significance	Odds ratio [Exp (β)]
Income:		14.38	$p<.05$	
Less than $20,000	1.12	9.20	$p<.01$	3.1
$20,000 to $29,999	0.56	2.64	$p<.10$	1.8
$30,000 to $49,999	0.93	9.51	$p<.01$	2.5
$50,000 to $74,999	0.65	4.50	$p<.05$	1.9
$75,000 to $99,999	0.40	1.80	n.s.	1.5
$100,000 to $149,999	0.61	3.57	p<.10	1.8
$150,000 or more[a]				
Education:		21.38	$p<.01$	
High School graduate	1.17	20.71	$p<.01$	3.2
Some college or voc. school	0.69	9.42	$p<.01$	2.0
College graduate	0.47	4.37	$p<.05$	1.6
Post graduate or adv. degree[a]				
Age:		10.14	$p<.10$	
18–24 years	−0.60	3.62	$p<.10$	0.5
25–34 years	0.21	0.59	n.s	1.2
35–44 years	0.33	1.60	n.s.	1.4
45–54 years	−0.04	0.03	n.s.	1.0
55–64 years	0.08	0.11	n.s.	1.1
65+ years[a]				
Gender:				
Male	0.47	4.42	$p<.05$	1.6
Female[a]				
Goodness-of-fit statistics:				
−2 Log Likelihood	1,082.38			
Model χ^2 (df=17)	191.05			
Significance	$p<.01$			
Cox and Snell R^2	0.19			
Nagelkerke R^2	0.25			

[a] Used as reference categories for β estimates

income category, which is consistent with the hypotheses. Similarly, an examination of the Odds Ratio coefficients for *Education* showed that respondents in the high school graduate category were also roughly three times more likely [(Exp (β)=3.2] to *Pay for Online Content* than those with a post graduate degree, which is also consistent with the hypotheses. An inspection of the magnitudes of the partial β's for *Income* and *Education* showed that willingness to pay for online content decreased with increasing levels of income and education as suggested by the hypotheses.

An examination of the Odds Ratio coefficients for *Gender* showed that males were approximately one and one-half times more likely [(Exp (β)=1.6] to *Pay for Online Content* than females, as predicted by the hypotheses. A review of the Odds Ratio coefficients for *Age* showed that respondents in the 25–34 years and 35–44 years age categories were 20 % and 40 % more likely, respectively, [(Exp (β)=1.2 and (Exp (β)= 1.4] to *Pay for Online Content* than respondents in the 65+years category, which is also supports the hypotheses. However, the Odds Ratio coefficients for *Age* need to be interpreted with caution because the overall relationship between *Age* and *Pay for Online Content* was only marginally significant (Wald's statistic=10.14; $p<.10$). Hence, it appears that *Gender* (female) is the primary driver of *Pay for Online Content* while *Age* has a secondary effect at best.

Next, the information theory-based artificial intelligence algorithm C5.0 (Quinlan 1992; Larose 2005) was used to validate the results obtained from the logistic regression analysis. An information gain (i.e., entropy reduction) measure was used to partition the data. The main advantage of the C5.0 classification model is that makes no statistical assumptions about the distribution of the variables used in the estimation. More importantly, the C5.0 algorithm assumes the effect of a variable in a subset of observations is unrelated to the effect of the same variable in other subsets of observations, thereby eliminating the need to explicitly specify moderating effects and/or interactions. Another key advantage of the C5.0 algorithm is that it produces "rule sets" (i.e., if–then statements) that are easier to interpret and implement by managers. Also, the decision maker has more leeway in selecting which rule-sets to implement and which to ignore, because the algorithm does *not* produce mutually exclusive rule-sets.

To estimate the C5.0 classification model, the dependent variables *Pay for Online Content* and *Amount Paid for Online Content* were simultaneously associated with the predictor variables, *Income*, *Education*, *Age*, and *Gender*, to generate rule-sets (i.e., association rules) that could be used to identify the characteristics of consumers are more likely to pay for online content. The specific rule-sets (i.e., association rules) as determined by the C5.0 algorithm that illustrate these demographic differences are shown in Table 4. For example, younger females with lower incomes are more likely to pay for online content (confidence=0.83), while older males with higher incomes are less likely to do the same (confidence=0.92). Taken together the 6 rule-sets depicted in Table 4 that describe the demographic profiles of consumers most willing to pay for online content confirm the results of the logistic regression analysis. Once again, *Gender* and *Age* emerge as the main determinants of *Pay for Online Content*, while *Income* and *Education* only have secondary effects.

Next, *Amount Paid for Online Content* was used as the dependent variable in a general linear model, while the demographic factors *Income*, *Education*, *Age*, and *Gender* were entered as independent variables. The regression model indicated a significant fit ($F=16.00$; 5 *df*; $p<.01$) with an Adjusted $R^2=.10$. Consistent with expectations, *Education* ($\beta=.14$; $t=3.75$; $p<.01$), *Income* ($\beta=.26$; $t=6.45$; $p<.01$), and *Gender* (male; $\beta=.12$; $t=3.01$; $p<.01$), were found to be positively related to *Amount Paid for Online Content*, as predicted by the hypotheses. Unfortunately, the expected relationship between *Age* and *Amount Paid for Online Content* failed to reach statistical significance ($\beta=-.04$ t$=-0.94$; n.s.). Thus, it seems that *Income*, *Education*, and *Gender* (male) are the main determinants of *Amount Paid for Online Content*, while *Age* does not have an effect.

Table 4 Rule sets for pay for online content by demographic segments

	Rule set	Rule confidence*
Pay for online content? (1=yes)	If Gender=female	c=0.83
	& Age=35 to 44 years	
	& Education=college graduate	
	& Income=$30,000 to $75,000	
	If Gender=female	c=0.75
	& Age=25 to 34 years	
	& Education=some college	
	& Income=$75,000 to $100,000	
	If Gender=male	c=0.75
	& Age=65+ years	
	& Education=college graduate	
	& Income=$50,000 to $75,000	
Pay for online content? (2=no)	If Gender=male	c=0.92
	& Age=55+ years	
	& Education=college graduate	
	& Income=$75,000 to $150,000	
	If Gender=male	c=0. 86
	& Age=35–54 years	
	& Education=post graduate degree	
	& Income≤$75,000	c=0.80
	If Gender=male	
	& Age=25 to 44 years	
	& Education=college graduate	
	& Income≥$150,000	

Note: *denotes proportion of respondents meeting rule set conditions that were correctly classified by the rule set

8 Findings

A comparison of those consumers who are most willing to pay for content and those who are not shows definitive contrasts in terms of gender and age and to a lesser degree in terms of income and education. Many of these relationships reverse, when the amount consumers may be willing to pay is factored into the mix. For instance, the results show that while females are more willing to pay for online content than males, the estimated amount consumers are likely to pay is more for males. Likewise, while consumers with higher-income and more education express a lower willingness to pay, the estimated amount they are likely to pay is higher in comparison to consumers with lower-income and less education.

An important finding is that there appears to be a “demographic divide” between consumers who are more likely to pay for online content and the amount they are likely to pay. Specifically, while willingness to pay for online content is related to age and gender, the amount paid for online content is more related to education and income. In other words, the ability to pay, in and of itself, does not translate into willingness to pay. This result is particularly important for online content providers using willingness to pay measures to assess the viability of a pay-for-content business model.

9 Limitations

The study was based data collected by a phone survey rather than online. Despite this limitation, the study is high in external validity because it is based on the real-world behavior of a nationally representative sample of 755 American internet users in 2010, within a sampling error of ±3.7 percentage points. To achieve the high degree of external validity some compromises had to be made during the data collection process. Several of the variables were measured using ordinal scales because of the concern that respondent fatigue might cause to prematurely terminate the phone interview, which would seriously affect sample representativeness.

10 Summary and conclusions

The empirical findings has important implications for an online content provider considering a transition from a content-for-free business model to a pay-for-content model, because they suggest that projected subscription revenues for various demographic segments may not necessarily align with the willingness to pay for online content reported by these segments. Specifically, the findings show that consumers who are more likely to pay for online content do so in lesser amounts, while somewhat ironically, those who are less likely to pay for online content do so in larger amounts. This somewhat counter-intuitive result has important implications for online content providers considering a change from an advertising-based revenue model to one that is subscription-based.

Why are some consumers less willing to pay for online content? A possible explanation is that online content providers who have long used an advertising supported content-for-free business model have created a “reference price of zero.” In other words, some consumers view any price above zero as a loss instead of a forgone gain. Such a prospect theory explanation is consistent with the data. The predictions from the economic theories of information search and product bundling, which suggest that search costs and bundling costs determine the amount consumers are likely to pay for online content are upheld by the current data. Specifically, the opportunity cost of time (i.e., economic time costs) does seem to have an effect on the amount consumers are likely to pay for online content. At the same time, communicating the information value of online content is likely to provide greater success in enhancing consumer willingness to pay, rather than tactics intended to limit access to content.

Acknowledgments The author gratefully acknowledges the financial support provided by the Connecticut Information Technology Institute (CITI) and thanks the Pew Internet & American Life Project for providing the data for the study.

References

Bourreau, M., & Lethiais, V. (2005). Pricing Information Goods: Free vs. pay content. In E. Brousseau & N. Curien (Eds.), *Internet and digital economics* (pp. 345–67). Cambridge, UK: Cambridge University Press.

Buchanan, J. M. (1965). An economic theory of clubs. *Economica, 32*(125), 1–14.

Chyi, H. I., & Yang, M. J. (2009). Is online news an inferior good? Examining the economic nature of online news among users. *Journalism and Mass Communication Quarterly, 86*(3), 594–612.

Clemons, E. K. (2009). Business models for monetizing internet applications and web sites: Experience, theory and predictions. *Journal of Management Information Systems, 26*(2), 15–41.

Gefen, D., & Ridings, C. M. (2005). If you spoke as she does, sir, instead of the way you do: A sociolinguistics perspective of gender differences in virtual communities. *Advances in Information Systems, 36*(2), 78–92.

Jackson, L. A., Ervin, K. S., Gardner, P. D., & Schmitt, N. (2001). Gender and the Internet: Women communicating and men searching. *Sex Roles, 44*(5/6), 363–379.

Katz, M. L., & Rosen, H. S. (1991). *Microeconomics*. Homewood, IL: Irwin.

Larose, D. T. (2005). *Discovering knowledge in data*. Hoboken, NJ: John Wiley & Sons.

Pauwels, K., & Weiss, A. (2008). Moving from free to fee: How online firms market to change their business model successfully. *Journal of Marketing, 72*, 14–31.

Quinlan, J. R. (1992). *C4.5: Programs for machine learning*, Morgan Kaufmann, San Francisco, CA.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy, 69*, 213–225.

Van Slyke, C., Comunale, C. L., & Belanger, F. (2002). Gender differences in perceptions of web-based shopping. Communications of the ACM, 45(8).