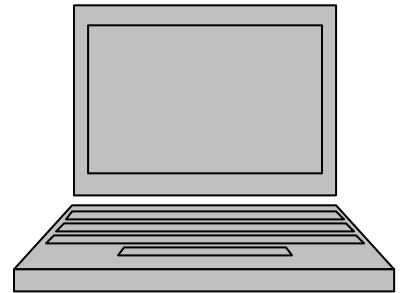
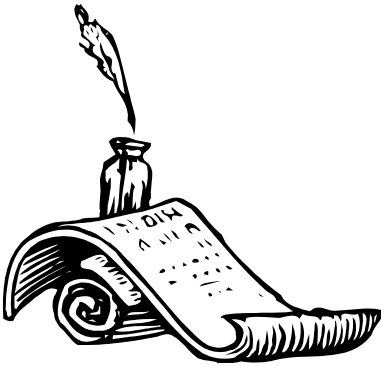


Assessing the Digital Divide:
An extension of the Pulse of the Community Survey



Terrence Willett
Director of Research
Gavilan Community College
Q&A Research, Inc.
March 2004

Introduction

In January 2004, Gavilan College completed a community survey with assistance from Q&A Research. The survey was funded by a Title V grant focused in part on closing the “digital divide” between Hispanics and non-Hispanics. The survey assessed community attitudes about the college and its direction as well as computer and internet access and skills. The complete survey consisted of 6 focus groups and 600 phone interviews in both English and Spanish from a representative cross section of the local community. Full details on the phone survey protocol including survey questions can be found in the Pulse of the Community report (www.gavilan.edu/research/reports/comm_phone_report).

The Pulse of the Community report found that Hispanics were significantly less likely than non-Hispanics to have access to computers (44% versus 79% respectively, Chi-square (1) = 77.977, $p < 0.0005$, $\phi = 0.361$). However, other demographic variables that are associated with ethnicity may help more fully illustrate the factors involved in computer access. This sub-report focuses examines the multivariate relationship between computer access and ethnicity, gender, age, language, income, and education.

Methods

Both a binary logistic regression and classification and regression tree (CART) provided an examination of these variables to eliminate method bias in the results. Computer access was dichotomously coded as either yes or no to the question of access. Ethnicity was dichotomously coded as Hispanic or non-Hispanic as the frequencies of non-Hispanic, non-White were too low for more detailed analysis. Gender was also dichotomously coded. Age was coded with six ordinal levels. Language was coded with five ordinal levels ranging as English only, mostly English, English and Spanish equally, mostly Spanish, and Spanish only with higher values indicating more Spanish use. Non-Hispanics were assumed to be in the English only category. Income was coded with seven ordinal levels. Education was coded with six ordinal levels.

The logistic regression was conducted with 11 excessively influencing cases removed as well as 120 cases with missing values leaving 469 cases for analysis. The model used a forward stepwise likelihood ratio model. Collinearity did not appear in the data. Residuals appeared approximately normal and a Hosmer and Lemeshow goodness-of-fit test indicated the final model fit the observed data reasonably well (Chi-square (8) = 4.523, $p = 0.807$).

The CART analysis used 598 cases after omitting two cases with missing data for computer access. A ten fold cross validation was used to test the pruned tree. The Gini impurity measure was used for splitting and pruning. Parent nodes were set to have a minimum of 50 cases while child nodes were set to have a minimum of 30 cases.

Results and Discussion

The logistic regression model included education, income, and language as predictors of computer access (Model Chi-square (15) = 209.192, $p < 0.0005$, Cox and Snell $R^2 = 0.361$, Nagelkerke $R^2 = 0.491$, risk = .208). Model parameters suggest that computer access increases with increasing education, income, and English use (Table 1).

The CART model also included education income, and language as predictors of computer access (risk = 0.246). The associations generally follow those of the logistic regression with language being a factor only for those with lower education levels and moderate to high incomes (Figure 1). The agreement between these two methods strengthens the case that these findings are relatively free from method bias.

While the univariate analysis indicated that there is likely a digital divide based upon Hispanic status within the Gavilan College service area, the multivariate analysis suggests that differences in computer access can be better explained by differences primarily in education and income and to some extent by language. The appearance of Hispanic status in the univariate analysis is likely due to the association between Hispanic status, income, education, and language use in the sample. It should be noted that this report did not examine causality between these variables but it could be assumed that some of these variables are “co-causal”. For example, those with higher incomes and education may more likely to have a preference for using computers and the use of computers in many cases can be a prerequisite for attaining higher education and income levels. This data is limited by the restricted geographic range of the sampling so that these results do not necessarily apply outside of the Gavilan College service area. In addition, all data are self-reported by survey respondents and have not been independently verified.

In the interest of closing the digital divide, these findings accentuate the importance of Gavilan College in the community as a provider of education and computer technology and training. This implies that the college should ensure that student computing resources are sufficiently abundant and up to date.

Table 1. Logistic model parameters.

Variables in the Equation	B	S.E.	Wald	df	Sig.	Exp(B)
Some High School			19.170	5	0.002	
High School Graduate	0.287	0.372	0.596	1	0.440	1.333
Some College	1.657	0.460	12.951	1	0.000	5.243
Business or Technical School	1.709	0.623	7.521	1	0.006	5.522
College Graduate	0.934	0.447	4.373	1	0.037	2.545
Postgraduate Degree	20.747	5461.519	0.000	1	0.997	1.024E+09
Under \$20k			34.739	6	0.000	
\$20k to < \$35k	-0.587	0.441	1.772	1	0.183	0.556
\$35k to < \$50k	0.801	0.455	3.093	1	0.079	2.227
\$50k to < \$75k	0.931	0.514	3.277	1	0.070	2.537
\$75k to < \$100k	0.580	0.486	1.424	1	0.233	1.786
\$100k to < \$150k	2.578	0.736	12.253	1	0.000	13.166
\$150k or more	20.334	6703.784	0.000	1	0.998	6.773E+08
English Only			9.570	4	0.048	
Mostly English	0.295	0.690	0.183	1	0.669	1.344
Both Spanish and English Equally	0.198	0.485	0.167	1	0.683	1.219
Mostly Spanish	-1.103	0.399	7.640	1	0.006	0.332
Spanish Only	-0.082	0.380	0.047	1	0.829	0.921
Constant	-0.798	0.470	2.884	1	0.089	0.450

