

***Technical Supplement***  
**to**  
***Social Effects of Culture:***  
***Detailed Statistical Models***



*Statistical Insights on the Arts, Vol. 7 No. 1*  
© Hill Strategies Research Inc., July 2008

**ABSTRACT**

Data drawn from Statistic Canada's Social Survey of 2005 is used to determine which factors, demographic and otherwise, can predict potential indicators of social engagement. Canadians 15 years of age or older were asked about their social engagement over various time periods: the past 12 months, the past month and currently. Six social indicators are examined as the independent variables in this report: doing unpaid volunteer work for an organization; donating money or goods to an organization or charity; doing a favour for a neighbour; feeling trapped in a daily routine; and sense of belonging to one's province and to Canada. Logistic regression analysis is utilized to determine the odds ratio for each of the independent variables included in the study. Model selection, variable selection, variable coding, model validity and results are discussed in this technical supplement.

## INTRODUCTION

The goal of this study is to examine demographic and cultural factors in selected social behaviour. Binary logistic regression is used to identify the best predictors of six social indicators: volunteering, donating, doing a favour for a neighbour, feeling trapped in a daily routine, sense of belonging to one's province and sense of belonging to Canada. The six social indicators are analyzed separately. That is, a different model is constructed where each of the six social indicators is represented as the dependent variable.

The data is drawn from Statistics Canada's General Social Survey of 2005, an in-depth telephone survey of about 20,000 Canadians 15 years of age or older. Because of a split in the survey design, only half of the respondents were asked questions about their cultural activities in 2005. A total of 9,851 respondents answered the cultural questions. The other half of the respondents were asked questions about social networking and trust, including some questions that, in theory, would be helpful in examining the social effects of culture. However, because of the survey split, these questions could not be cross-tabulated with the questions about cultural activities.

## METHODS

### Model Selection

Logistic regression is used to predict the probability of an event occurring due to the occurrence of other events. More specifically, binary logistic regression is used when the single dependent variable (the event that is being predicted) is dichotomous and the independent variables (other events that have already occurred) are of any type (binary, continuous, categorical, etc.). This method is used in order to categorize predictor variables based on relative significance, to gauge multicollinearity of the independent variables and to understand the effect of the predictor variables on the dependent variable.

Binary (or binomial) logistic regression is a form of analysis where logistic regression equations are solved iteratively, a contrast from other forms of regression analysis where a mathematical equation is solved explicitly. In binary logistic regression a trial equation is fitted and adjusted in order to improve the fit:

$$\text{logit}[p] = \log [p/1-p] = \alpha + \beta_1x_1 + \beta_2x_2 + \dots + \beta_ix_i \quad [\text{Equation 1}]$$

where  $p$  is the probability of the event occurring,  $\alpha$  is the constant of the equation and  $\beta_i$  are the coefficients of the predictor values. The user determines when the iterations stop. That is, either an explicit number of iterations are performed or a cut-value is selected so that the iterations stop when the improvement from step to the next is less than the cut-value.

Once the equation is fitted one can determine the odds ratio. The odds ratio is a prediction about how likely the respondent is to have participated in the activity in question. This is calculated by raising the base of the natural logarithm to the  $\beta$ th power, where  $\beta$  is the coefficient of the predictor variable from the trial equation. That is:

$$\text{ODDS} = e^{\beta} = p/1-p \quad [\text{Equation 2}]$$

where  $\beta$  is the slope of the logistic regression equation for a specific predictor variable. Once the odds of an event occurring have been calculated, they can be converted to probabilities:

$$p = \text{ODDS} / (1 + \text{ODDS}) \quad [\text{Equation 3}]$$

Probabilities can be more useful when describing the effects of the predictor variables on the dependent variable.

Predictive analytic software, SPSS, is used to perform the iterations necessary to derive the logistic regression equation. Once the equation is fitted, the output presents the user with the coefficients of the predictor variables and also includes a 95% confidence interval used for error analysis. SPSS also has intrinsic tests to examine the validity of the created models. These tests are discussed more thoroughly in a subsequent section.

### **Variable Selection**

The predictor variables can be divided into two categories: demographic variables and cultural variables. For the cultural variables, a cultural participant is defined as someone who participated at least once in the relevant cultural activity in 2005. An important difference from a previous report that examined cultural factors in social engagement is that the general performing arts attendance category has been separated into three component parts (theatre attendance, pop concert attendance and classical music attendance).

Preliminary exploration is done to discover trends in the dataset for each of the six social indicators under investigation. Specifically, the preliminary exploration examines the percentage of respondents who have volunteered (for example), given their demographic characteristics and cultural activities.

Frequency distributions are constructed to examine the percentage of respondents in each category of the social indicators (the dependent variables) based on their demographic characteristics and cultural activities (the independent variables). Using frequency distributions, the percentage of respondents in each independent variable having the social characteristic (or behaviour) is calculated. Both positive and negative effects are examined. The results are evaluated and a subset is constructed of the independent variables most likely to predict social behaviour for each of the six dependent variables. The independent variables with the highest and lowest percentages are considered as potential primary predictors of the social behaviour.

Twenty-five variables are selected in total plus an additional reconfiguration variable representing the respondents' ages. Of the 25 chosen variables, 14 are demographic and 11 are cultural. Brief explanations of each variable are provided in Appendix 2.

Once the base set of variables to be included in the study is constructed, exploratory tests are done to discover which predictor variables have large influences on each of the dependent variables. This is done by examining the differences between the percentages of those who have

participated and those who have not. For example, 82% of readers indicated that they had donated in the past 12 months, while only 66% of non-readers had donated in the past 12 months, resulting in a considerable difference of 16%.

A model is constructed for each dependent variable using the enter procedure for a logistic regression. Initially, the variables deemed to be the most influential are included. The remaining variables are tested in the model and are kept if they increase the predictive accuracy.

Three main statistical tests are used to test the validity of the models: -2 log likelihood, the Hosmer and Lemeshow test and tests for possible multicollinearity. These methods will be discussed in the section “Assessment of Study Quality”.

The variables that the preliminary research indicates may be of primary importance are included in the initial drafts of the models, while the remaining variables are added and removed in order to improve the accuracy of the model. This is done without preference for any of the variables. For example, if adding a variable does not cause the -2 log likelihood to change significantly, then the variable in question does not have much importance in predicting the dependent variable and can be left out of the model. These steps of adding and removing different combinations of variables is done until the -2 log likelihood is as small as possible, while the Hosmer and Lemeshow test show that the data fit the model well and the tests for multicollinearity do not show multicollinearity of the predictor variables.

### **Variable Coding**

A necessity of a binary logistic regression is that the dependent variable is dichotomous. Consequently, all six of the social indicators have been transformed to satisfy the condition. Many of the variables did not require significant transformations into a dichotomous variable. For example, volunteering had four available responses: yes, no, don't know, no response. The variable was transformed to only retain two responses, yes and no. This process also applies to the donating and feeling trapped in a daily routine variables. Other variables, such as one's sense of belonging (to one's province and to Canada) were more complicated to transform. For these variables, there are seven available responses: very strong, somewhat strong, somewhat weak, very weak, no opinion, don't know and no response. These variables were transformed in order to predict a very strong sense of belonging. That is, the “somewhat strong”, “somewhat weak” and “very weak” were coded to the same position and the “no opinion”, “don't know” and “no response” categories were not retained. This leaves only two responses: very strong and other (i.e., somewhat strong, somewhat weak or very weak).

The 25 base variables and additional reconfigured variable are included in the model as independent variables with the internal value “0” being set as the reference level. More specifically, each variable has internal numerical coding that reflects a specific answer. For example, variables that have two possible elements in the answer set are labeled dichotomous. These are fairly straightforward to analyze, as their coding would result in one element being the reference level and the other being compared to it. For example, if the respondent answered “no” to question LCA\_Q330, then they did not attend a symphonic or classical music performance during the past 12 months. This result is classified as the reference variable with coding “0”. The

only other available response would be that the respondent had attended a symphonic or classical music performance during the past 12 months (coded “1”).

For most of the dichotomous variables, “no” was coded “0” and “yes” was coded “1”. As a consequence, the results show that there is a direct effect on the dependent variable. For example, if the odds given in the output is 1.472 (Table 1: Results of the volunteering model) then the implication is that someone who has attended a symphonic or classical music performance during the past 12 months is 1.472 times as likely to have also done unpaid volunteer work for any organization during the past 12 months.

Many of the questions asked by Statistics Canada include the options of “don’t know”, “not asked” and “not stated”. For example, the question regarding the country of birth of the respondent has four possible responses: “Canada”, “country outside of Canada”, “not stated” and “don’t know”. The responses “not stated” and “don’t know” are not useful in predicting if the respondent has a very strong sense of belonging to Canada (for example). For this reason, the independent variable has been altered to become a dichotomous variable. This variable is coded as “0” if Canada was the country of birth of the respondent and “1” for a country outside of Canada.

Another example of this is the question regarding the household language of the respondent. The respondent’s possible answers are English only, French only, other language, not stated and don’t know. In this case, when the two responses “not stated” and “don’t know” are removed, a categorical variable remains. Coding of the variables in the categorical case is more complicated than for dichotomous variables. The question regarding the household language of the respondent is coded in the following manner: English only = 0, French only = 1, other language = 2. In this case, English represents the reference point, and each of the other two possible responses would display an influence on the dependent variable relative to the reference point.

Determining the importance of the elements of a categorical variable is not as simple as the dichotomous case. Looking at the SPSS output in Table 3: the favour for a neighbour model, there are three possible responses for the household language of the respondent. With English as the only language being the reference element, the two other scenarios display odds relative to it. Respondents with French as their only household language are .826 times as likely to have done a favour for a neighbour in the previous month as respondents speaking English in their household (from Table 3: Exp(B) column). Respondents with a household language other than English or French are .790 times as likely to have done a favour for a neighbour in the last month as respondents who speak English in their household (from Table 3: Exp(B) column).

But what if a comparison between those who speak French in their household and those who speak a language other than French or English in their household is necessary? Since the ratios are: 1:0.826 and 1:0.790 for French speaking households and households speaking a language other than English or French, respectively, one can compare the two directly; 0.826:0.790. To help with the ease of comparison one could divide each side by one of the two values (the value chosen being the new reference variable). Dividing each side by 0.790 shows that respondents speaking French in their household are 1.05 times more likely to have done a favour for a

neighbour during the past month than respondents speaking a language other than French or English in their household.

## RESULTS

### Assessment of Study Quality

The validity of the models was determined using three fundamental tests: the likelihood ratio test, the Hosmer and Lemeshow test and a test for multicollinearity.

The log likelihood procedure is used to determine the difference between a model and a nested model, which is a subset of the original model. If there is no significant difference between the log likelihood of the model and the nested model, then the variable is dropped to create a more concise model. This is used to determine which variables are necessary to improve the prediction of the dependent variable.

The Hosmer and Lemeshow test is used to determine if the model's approximations fit the data adequately. If the Hosmer and Lemeshow test statistic is not significant, that is, greater than .05, we satisfy the hypothesis that the actual and model predicted values are analogous. As a result, for the model to be good fitting, non-significant results on the Hosmer and Lemeshow test are required. This test is only an interim step in ensuring validity of the overall model.

Another problem that a model could have is multicollinearity. This occurs when there are independent variables in the model that have a strong correlation. The degree of the correlation is described by a correlation coefficient, a value between -1 and 1 that measures the degree of relationship between two values: 1 being completely related and -1 being completely oppositely related. A problem occurs when two or more variables have high correlation coefficients. In this case, the variance in the dependent variable due to some of the independents will not be legitimate. Consequently, as each of the models is analyzed, a matrix of correlation coefficients is constructed and each entry analyzed. If one element is greater than 0.90 or if several are greater than 0.70, multicollinearity could be a problem and the model should be discarded. None of the independent variables in the models described in this report have strong multicollinearity.

Exp(B) is the odds ratio for each independent variable and is a positive real number. If exp(B) is less than one, then the independent variable predicts a lower likelihood of engaging in the social behaviour than those in the reference group. If exp(B) is greater than one, it predicts a higher likelihood of engaging in the social behaviour. A useful error indicator of the value exp(B) is the 95% confidence interval. The 95% confidence interval is a range around the value exp(B) in the SPSS output. It provides a scope of uncertainty for the odds ratio for each independent variable. That is, it accounts for various types of error (sample, non-sampling) that could have an effect on the results. The 95% confidence interval helps to determine how useful each independent variable is as a predictor of the dependent variable. A problem arises when the 95% confidence interval includes both predictions: an exp(B) with both a value less than one and greater than one. If this occurs, the independent variable is not considered to be a useful predictor of the dependent variable.

## **Data Interpretation**

The meaningful results obtained from the SPSS output are provided in Appendix 1 for each of the six models (volunteering, donating, doing a favour for a neighbour, feeling trapped in a daily routine, sense of belonging to one's province and sense of belonging to Canada). In each of the tables, the value in the column labeled "Exp(B)" represents the odds ratio for each of the independent variables. For dichotomous variables, the odds ratio (exp(B)) indicates that someone who responded "yes" is exp(B) times as likely to have the given social characteristics than someone who responded "no" to the same question. For example, in Table 4, (the feeling trapped in a daily routine model) the "ACTLIMIT" variable is dichotomous. The reference level for this variable is the respondent not having any activity limitations. The exp(B) value in the table, 1.833, is the odds ratio. This indicates that someone who has activity limitations is 1.833 times as likely as someone with no limitations to feel trapped in a daily routine.

For categorical variables, if the respondent fits into one of the classes of the variable, then they are exp(B) times as likely to participate compared to someone in the reference class. The reference class is the first option listed and always has an empty entry for exp(B). For instance, the variable "AGEGR10" in Table 2 has 7 classes. The first variable listed, 15 to 24, is the reference level. The results are read as follows: the exp(B) for someone who is 75 or older is 6.246. Therefore, someone who is 75 or older is 6.246 times as likely to have donated money or goods to any organization or charity (not including membership fees or dues) in the past 12 months as someone who is 15 to 24.

Further, for each of the classes that has a value for exp(B), there are columns entitled "Lower" and "Upper". These represent the 95% confidence interval for the odds ratio, more specifically the error in the odds ratio. If the 95% confidence interval contains the value "1", then it contains both a prediction of "yes" and "no" with regards to the dependent variable. This corresponds to a value larger than .05 in the "sig." column indicating that the independent variable is not an effective predictor in the model.

All dichotomous variables that are not considered significant predictors are excluded from the models. Care must be used when examining the results for the categorical variables since not all classes are necessarily significant. Categorical variables with at least one class that is useful are included in the model. Looking at the results for the categorical variable MARSTAT in Table 1, there are several groupings that have values larger than .05 in the "sig." column. The classes are: widowed, divorced and single (never married). Therefore, the useful predictors in the model are the categories married, common-law and separated.

**APPENDIX 1: SPSS OUTPUT**

**Table 1: Results of the volunteering model (SPSS output for binary logistic regression model)**

	B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
							Lower	Upper
LCA_Q310- "did you attend a theatrical performance such as a drama, musical theatre, dinner theatre, comedy?" (Reference is not going)	.377	.075	25.514	1	.000	1.458	1.259	1.687
LCA_Q320- "did you attend a popular musical performance such as pop/rock, jazz, blues, folk, country and western?" (Reference is not going)	.153	.073	4.451	1	.035	1.166	1.011	1.345
LCA_Q330- "did you attend a symphonic or classical music performance?" (Reference is not going)	.387	.103	14.127	1	.000	1.472	1.203	1.801
LCA_Q340- "did you go to a cultural or artistic festival?" (Reference is not going)	.339	.073	21.387	1	.000	1.403	1.216	1.620
LCA_Q300- "did you go to a concert or performance by professional artists of music, dance, theatre, or opera, excluding cultural festivals?" (Reference is not going)	.266	.081	10.806	1	.001	1.305	1.113	1.529
LCA_Q410- "did you go to a public art gallery or art museum?" (Reference is not going)	.179	.071	6.396	1	.011	1.196	1.041	1.374
LCA_Q430- "did you go to an historic site?" (Reference is not going)	.425	.064	43.512	1	.000	1.529	1.348	1.735
LCA_Q450- "did you go to a conservation area or nature park?" (Reference is not going)	.138	.063	4.866	1	.027	1.148	1.016	1.299
LCA_Q110- "did you read a newspaper?" (Reference is not reading)	.275	.093	8.719	1	.003	1.317	1.097	1.581
LCA_Q130- "did you read a book?" (Reference is not reading)	.258	.067	14.633	1	.000	1.294	1.134	1.477



		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
EDU5*	College diploma or some college			37.023	2	.000			
	High school or less	-.306	.071	18.773	1	.000	.736	.641	.846
	Degree	.217	.074	8.551	1	.003	1.242	1.074	1.437
BRTHCAN- "country of birth of respondent" (Reference is birth in Canada)		-.232	.094	6.050	1	.014	.793	.659	.954
LUC_RST- "urban/rural indicator" (Reference is urban areas)		.467	.070	45.011	1	.000	1.595	1.392	1.828
INCMHSD*	Less than \$40,000			10.877	2	.004			
	\$40,000 to \$79,999	.197	.072	7.423	1	.006	1.217	1.057	1.402
	\$80,000 or more	.262	.085	9.466	1	.002	1.300	1.100	1.537
SEX- "sex of respondent" (Reference is female)		.136	.060	5.092	1	.024	1.145	1.018	1.289
RELIGATT	At least once a week			223.423	4	.000			
	At least once a month	-.499	.097	26.270	1	.000	.607	.502	.735
	A few times a year	-.850	.086	98.022	1	.000	.427	.361	.506
	At least once a year	- 1.194	.114	110.072	1	.000	.303	.242	.379
	Not at all	- 1.206	.087	191.968	1	.000	.300	.253	.355
CHRFLAG- "single child(ren) of the respondent living in the household" (Reference is having child(ren) in the household)		.253	.072	12.520	1	.000	1.288	1.120	1.482

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
MARSTAT (Reference is married)	Married			17.828	5	.003			
	Living common law	-.317	.113	7.930	1	.005	.728	.584	.908
	Widowed	.104	.127	.668	1	.414	1.109	.865	1.422
	Separated	.370	.148	6.255	1	.012	1.448	1.083	1.936
	Divorced	-.105	.117	.808	1	.369	.900	.716	1.132
	Single (never married)	-.050	.092	.292	1	.589	.952	.795	1.139
AGEGR10* (Reference is 15 to 24)	15 to 24			21.877	3	.000			
	25 to 44	-.489	.113	18.804	1	.000	.613	.492	.765
	45 to 64	-.532	.118	20.451	1	.000	.587	.466	.740
	65 or older	-.480	.143	11.315	1	.001	.619	.468	.818
RELIG6 (Reference is no religion)	No religion			32.494	5	.000			
	Roman Catholic	-.328	.101	10.626	1	.001	.720	.591	.877
	United Church	.046	.119	.148	1	.701	1.047	.829	1.321
	Protestant	.079	.100	.615	1	.433	1.082	.889	1.317
	Other	-.141	.151	.869	1	.351	.869	.646	1.168
	Para-religious groups or unknown	-.500	.608	.676	1	.411	.607	.184	1.998
LANHSDC (Reference is English)	English only			17.695	2	.000			
	French only	-.260	.089	8.600	1	.003	.771	.648	.917
	Other language	-.432	.133	10.572	1	.001	.649	.500	.842
Constant		-.400	.198	4.060	1	.044	.671		

**Table 2: Results of the donating model (SPSS output for binary logistic regression model)**

	B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)		
							Lower	Upper	
LCA_Q110- "did you read a newspaper?" (Reference is not reading)	.496	.092	29.338	1	.000	1.642	1.372	1.965	
LCA_Q130- "did you read a book?" (Reference is not reading)	.470	.073	41.185	1	.000	1.600	1.386	1.848	
LCA_Q210- "did you go to a movie or drive-in?" (Reference is not going)	.354	.080	19.561	1	.000	1.425	1.218	1.667	
LCA_Q310- "did you attend a theatrical performance such as a drama, musical theatre, dinner theatre, comedy?" (Reference is not going)	.391	.105	13.892	1	.000	1.478	1.204	1.816	
LCA_Q430- "did you go to an historic site?" (Reference is not going)	.561	.089	40.193	1	.000	1.753	1.474	2.085	
LCA_Q450- "did you go to a conservation area or nature park?" (Reference is not going)	.482	.079	37.575	1	.000	1.620	1.388	1.890	
AGEGR10	15 to 24		167.728	6	.000				
	25 to 34	.272	.126	4.623	1	.032	1.312	1.024	1.681
	35 to 44	.808	.129	39.032	1	.000	2.244	1.741	2.891
	45 to 54	.801	.127	39.584	1	.000	2.227	1.735	2.858
	55 to 64	1.131	.138	67.330	1	.000	3.099	2.365	4.060
	65 to 74	1.587	.162	95.612	1	.000	4.891	3.558	6.722
	75 years and over	1.832	.190	92.920	1	.000	6.246	4.304	9.065
INCMHSD*	Less than \$40,000		105.393	2	.000				
	\$40,000 to \$79,999	.659	.081	65.503	1	.000	1.933	1.648	2.268
	\$80,000 or more	.930	.103	81.202	1	.000	2.535	2.070	3.103

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
RELIGATT	At least once a week			57.308	4	.000			
	At least once a month	-.390	.139	7.834	1	.005	.677	.515	.890
	A few times a year	-.351	.128	7.499	1	.006	.704	.548	.905
	At least once a year	-.823	.146	31.765	1	.000	.439	.330	.585
	Not at all	-.777	.119	42.545	1	.000	.460	.364	.581
CHRFLAG- "single child(ren) of the respondent living in the household"		.219	.084	6.753	1	.009	1.245	1.055	1.469
EDU5*	College diploma or some college			33.544	2	.000			
	High school or less	-.392	.079	24.598	1	.000	.676	.579	.789
	Degree	.140	.106	1.746	1	.186	1.150	.935	1.414
REGION	Ontario			43.832	4	.000			
	Atlantic region	.447	.108	17.247	1	.000	1.563	1.266	1.930
	Quebec	.333	.105	10.134	1	.001	1.395	1.137	1.713
	Prairie region	-.092	.098	.880	1	.348	.912	.753	1.105
	British Columbia	-.211	.118	3.183	1	.074	.810	.642	1.021
RLR_Q105*	Religion is somewhat or not very important			14.373	2	.001			
	Religion is very important	.188	.086	4.746	1	.029	1.207	1.019	1.429
	Religion is not very important	-.373	.133	7.811	1	.005	.689	.530	.895
Constant		-.654	.190	11.865	1	.001	.520		

**Table 3: Results of the favour for a neighbour model (SPSS output for binary logistic regression model)**

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
SEX- "sex of respondent" (Reference is female)		-.175	.057	9.376	1	.002	.840	.751	.939
INCMHSD*	Less than \$40,000			16.225	2	.000			
	\$40,000 to \$79,999	.135	.065	4.334	1	.037	1.145	1.008	1.300
	\$80,000 or more	.317	.079	16.188	1	.000	1.373	1.176	1.602
CHRFLAG- "single child(ren) of the respondent living in the household"		.215	.061	12.234	1	.000	1.239	1.099	1.398
LANHSDC	English only			10.798	2	.005			
	French only	-.192	.070	7.550	1	.006	.826	.720	.947
	Other language	-.235	.117	4.072	1	.044	.790	.629	.993
LCA_Q110- "did you read a newspaper?" (Reference is not reading)		.343	.078	19.171	1	.000	1.410	1.209	1.644
LCA_Q130- "did you read a book?" (Reference is not reading)		.143	.063	5.195	1	.023	1.154	1.020	1.305
LCA_Q300- "did you go to a concert or performance by professional artists of music, dance, theatre, or opera, excluding cultural festivals?" (Reference is not going)		.292	.083	12.519	1	.000	1.339	1.139	1.575
LCA_Q410- "did you go to a public art gallery or art museum?" (Reference is not going)		.227	.071	10.289	1	.001	1.255	1.092	1.441
LCA_Q430- "did you go to an historic site?" (Reference is not going)		.167	.066	6.469	1	.011	1.182	1.039	1.344
LCA_Q450- "did you go to a conservation area or nature park?" (Reference is not going)		.321	.061	27.416	1	.000	1.378	1.222	1.554

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
EDU5*	College diploma or some college			12.836	2	.002			
	High school or less	-.156	.066	5.594	1	.018	.856	.752	.974
	Degree	-.243	.074	10.911	1	.001	.784	.678	.906
LUC_RST- "urban/rural indicator" (Reference is urban areas)		.266	.069	14.779	1	.000	1.305	1.139	1.495
BRTHCAN- "country of birth of respondent" (Reference is birth in Canada)		-.239	.087	7.625	1	.006	.787	.664	.933
LCA_Q210- "did you go to a movie or drive-in?" (Reference is not going)		-.156	.061	6.497	1	.011	.856	.759	.965
Constant		.377	.108	12.085	1	.001	1.458		

**Table 4: Results of the feeling trapped in a daily routine model (SPSS output for binary logistic regression model)**

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
SEX- "sex of respondent" (Reference is female)		.301	.049	38.195	1	.000	1.351	1.228	1.486
RELIGATT	At least once a week			16.826	4	.002			
	At least once a month	.012	.084	.020	1	.886	1.012	.859	1.193
	A few times a year	.156	.072	4.660	1	.031	1.169	1.014	1.346
	At least once a year	.224	.089	6.281	1	.012	1.251	1.050	1.491
	Not at all	.230	.067	11.714	1	.001	1.259	1.103	1.436
CHRFLAG- "single child(ren) of the respondent living in the household"		.334	.054	37.744	1	.000	1.397	1.256	1.554
LANHSDC	English only			62.255	2	.000			
	French only	.460	.060	59.201	1	.000	1.583	1.408	1.780
	Other language	.244	.087	7.935	1	.005	1.276	1.077	1.512

	B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
							Lower	Upper
LCA_Q310- "did you attend a theatrical performance such as a drama, musical theatre, dinner theatre, comedy?" (Reference is not going)	-.139	.063	4.875	1	.027	.870	.769	.984
LCA_Q330- "did you attend a symphonic or classical music performance?" (Reference is not going)	-.305	.093	10.649	1	.001	.737	.614	.885
LCA_Q410- "did you go to a public art gallery or art museum?" (Reference is not going)	-.206	.058	12.405	1	.000	.814	.726	.913
ACTLIMIT- "respondent is limited in the amount of physical activity that he/she can do" (Reference is not being limited)	.606	.068	80.365	1	.000	1.833	1.606	2.093
LCA_Q130- "did you read a book?" (Reference is not reading)	-.116	.052	4.925	1	.026	.890	.804	.987
AGEGR10*	15 to 24		267.958	3	.000			
	25 to 44	-.001	.079	.000	1	.990	.999	1.166
	45 to 64	-.340	.077	19.692	1	.000	.712	.827
	65 or older	-1.341	.097	189.722	1	.000	.262	.317
Constant	-.730	.094	60.227	1	.000	.482		

**Table 5: Results of the sense of belonging to province model (SPSS output for binary logistic regression model)**

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
EDU5*	College diploma or some college			7.912	2	.019			
	High school or less	.130	.055	5.591	1	.018	1.139	1.023	1.269
	Degree	-.040	.066	.375	1	.540	.960	.844	1.093
REGION	Ontario			60.846	4	.000			
	Atlantic region	.515	.071	52.286	1	.000	1.674	1.456	1.925
	Quebec	.440	.077	32.411	1	.000	1.553	1.334	1.807
	Prairie region	.341	.071	22.858	1	.000	1.406	1.223	1.616
	British Columbia	.305	.085	12.978	1	.000	1.356	1.149	1.601
RLR_Q105*	Religion is somewhat or not very important			78.529	2	.000			
	Religion is very important	.457	.052	77.991	1	.000	1.580	1.428	1.749
	Religion is not very important	.242	.084	8.277	1	.004	1.273	1.080	1.501
LCA_Q410- "did you go to a public art gallery or art museum?" (Reference is not going)		.192	.059	10.589	1	.001	1.212	1.080	1.361
LCA_Q430- "did you go to an historic site?" (Reference is not going)		.180	.054	11.163	1	.001	1.198	1.077	1.332
Age	Age 15 to 44			135.493	2	.000			
	Age 45 to 74	.472	.053	78.125	1	.000	1.604	1.444	1.781
	Age 75 and older	.984	.094	109.736	1	.000	2.674	2.225	3.214
BRTHCAN- "country of birth of respondent" (Reference is birth in Canada)irth_location(1)		-.283	.075	14.280	1	.000	.754	.651	.873
RELIG6	No religion			36.623	5	.000			
	Roman Catholic	.402	.079	25.961	1	.000	1.495	1.281	1.744
	United Church	.533	.096	30.538	1	.000	1.704	1.411	2.059
	Protestant	.327	.082	16.111	1	.000	1.387	1.182	1.628
	Other	.384	.128	9.001	1	.003	1.468	1.142	1.887
	Para-religious group or unknown	.465	.482	.931	1	.335	1.592	.619	4.092
LCA_Q210- "did you go to a movie or drive-in?" (Reference is not going)		-.177	.054	10.985	1	.001	.837	.754	.930
Constant		-2.020	.099	416.589	1	.000	.133		



**Table 6: Results of the sense of belonging to Canada model (SPSS output for binary logistic regression model)**

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
BRTHCAN- "country of birth of respondent" (Reference is birth in Canada)		-.283	.071	15.616	1	.000	.754	.655	.867
LCA_Q310- "did you attend a theatrical performance such as a drama, musical theatre, dinner theatre, comedy?" (Reference is not going)		.150	.057	6.786	1	.009	1.161	1.038	1.300
LCA_Q410- "did you go to a public art gallery or art museum?" (Reference is not going)		.188	.056	11.164	1	.001	1.206	1.081	1.347
LCA_Q430- "did you go to an historic site?" (Reference is not going)		.125	.053	5.698	1	.017	1.134	1.023	1.257
LCA_Q450- "did you go to a conservation area or nature park?" (Reference is not going)		.135	.049	7.428	1	.006	1.144	1.039	1.261
AGEGR10	15 to 24			105.106	6	.000			
	25 to 34	.211	.090	5.472	1	.019	1.235	1.035	1.474
	35 to 44	.332	.092	12.905	1	.000	1.393	1.163	1.669
	45 to 54	.400	.093	18.312	1	.000	1.492	1.242	1.792
	55 to 64	.642	.100	41.177	1	.000	1.901	1.562	2.313
	65 to 74	.920	.112	67.637	1	.000	2.509	2.015	3.124
	75 or older	1.108	.130	72.508	1	.000	3.027	2.346	3.907
CHRFLAG- "single child(ren) of the respondent living in the household"		-.182	.059	9.602	1	.002	.833	.742	.935
LANHSDC	English only			63.318	2	.000			
	French only	-.817	.113	52.413	1	.000	.442	.354	.551
	Other language	-.419	.097	18.707	1	.000	.658	.544	.795
RLR_Q105*	Religion is somewhat or not very important			81.404	2	.000			
	Religion is very important	.432	.049	78.817	1	.000	1.541	1.401	1.695
	Religion is not very important	.048	.072	.447	1	.504	1.050	.911	1.209

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
REGION	Ontario			48.518	4	.000			
	Atlantic region	.003	.065	.002	1	.966	1.003	.883	1.139
	Quebec	-.739	.114	41.930	1	.000	.478	.382	.597
	Prairie region	-.138	.063	4.777	1	.029	.871	.770	.986
	British Columbia	-.116	.074	2.455	1	.117	.891	.771	1.029
MARSTAT	Married			12.393	5	.030			
	Living common law	-.098	.090	1.205	1	.272	.906	.760	1.080
	Widowed	.053	.098	.287	1	.592	1.054	.870	1.277
	Separated	-.237	.119	3.978	1	.046	.789	.625	.996
	Divorced	-.042	.089	.225	1	.635	.959	.806	1.141
	Single (never married)	-.209	.070	8.841	1	.003	.812	.708	.931
Constant		-.332	.102	10.503	1	.001	.717		

**Table 5: Hosmer and Lemeshow significance values for the four models**

	Volunteering	Donating	Favour for a neighbour	Daily routine	Sense of belonging to province	Sense of belonging to Canada
Hosmer and Lemeshow Test (sig.)	0.986	0.169	0.580	0.841	.618	.078

## APPENDIX 2: VARIABLES INCLUDED IN STUDY

In all cases, the reference class is listed first.

\*indicates that two or more of the available answers for the specific question were grouped together for analysis purposes

† indicates that some of the available responses were excluded from analysis (most commonly the responses “don’t know”, “not stated” and “not asked”)

### DEMOGRAPHIC VARIABLES

#### Age

*Age is thought to be an important factor in social engagement. The variable was transformed several times in order to determine if some age groups were better predictors of the various dependent variables. This was done to distinguish groups that were not distributed evenly compared to others. For example, if the weighted percentage of individuals in some of the various classes of the variable were approximately equal, then the overall range of the age groups would be a better estimate than the smaller arrays.*

- Age group of respondent (AGEGR10, Position 48 in the GSS dataset)
  - 15 to 24
  - 25 to 34
  - 35 to 44
  - 45 to 54
  - 55 to 64
  - 65 to 74
  - 75 and older
- Age group of respondent (based on AGEGR10, Position 48)\*
  - 15 to 24
  - 25 to 44
  - 45 to 64
  - 65 or older

#### Sex

*Sex is thought to be an important factor in social engagement.*

- Sex of respondent (SEX, Position 50)
  - Female
  - Male

## **Marital Status**

*Does the relationship status of the respondent have an effect on how socially active they are? It could be that people who are married or in a common law relationship are less likely to participate in social activities as frequently as those who are single since single individuals may have more free time and thus less likely to feel trapped in a daily routine.*

- Marital status of respondent (MARSTAT, Position 51) †
  - Married
  - Living common-law
  - Widowed
  - Separated
  - Divorced
  - Single (Never married)

## **Single Children in the Household**

*Having a single child or children living in the household could introduce additional opportunities to socialize or restrict one's time to do so.*

- Single child(ren) of the respondent living in the household (CHRFLAG, Position 58)
  - Yes
  - No

## **Region**

*The region of residence of the respondent could be an important demographic factor, especially regarding the sense of belonging to one's province.*

- Region of residence of the respondent (REGION, Position 72)
  - Ontario
  - Atlantic region
  - Quebec
  - Prairie region
  - British Columbia

## **Urban/Rural Indicator**

*The size of the location in which the respondent lives could affect their social behaviour.*

- Urban/Rural indicator (LUC\_RST, Position 73) †
  - Larger Urban Centres (CMA/CA)
  - Rural and Small Town (non-CMA/CA)

## **Activity Limitations**

*Having activity limitations or disabilities could affect people's social engagement.*

- Respondent is limited in the amount of physical activity he/she can do at home, at work, at school or in other activities because of a physical condition or a mental condition or health problem (ACTLIMIT, Position 1640) †
  - No
  - Yes

## **Education**

*Education is thought to be an important factor in social engagement. The question regarding the amount of education acquired by the respondent was altered into three categories: no post-secondary education, some post secondary education and completion of a university degree or more.*

- Highest level of education obtained by the respondent (EDU5, Position 1737) † \*
  - Diploma/certificate from community college or trade/technical or some university/community college
  - High school diploma or some secondary/elementary/no schooling
  - Doctorate/masters/bachelor's degree

## **Country of Birth**

*This variable was included in the study because it is the best available indicator of whether early childhood cultural impacts could potentially affect the respondent in adulthood. Immigrants' sense of belonging to Canada is an important question.*

- Country of birth of respondent (BRTHCAN, Position 2222)†
  - Canada
  - Country outside Canada

## **Household Language**

*Language and official language minorities may have an impact on sense of belonging to their province and/or Canada.*

- Respondent's household language (LANHSDC)† Position:2239
  - English only
  - French only
  - Other language

## **Religion**

*Religion is a potential predictor of social engagement. Most sects of religion are based upon morals and consequently could affect an individual's likeliness of donating, volunteering and/or helping their neighbours.*

- Religion of respondent In six categories (RELIG6, Position 2240)†
  - No religion
  - Roman Catholic
  - United Church
  - Protestant
  - Other
  - Para-religious groups or unknown
- How important are your (religious or) spiritual beliefs to the way you live your life? Would you say they are: (based on RLR\_Q105 Position 2241)†\*
  - Very important
  - Somewhat important or not very important
  - Not at all important
- Religious attendance of the respondent (RELIGATT, Position 2242)†
  - At least once a week
  - At least once a month
  - A few times a year
  - At least once a year
  - Not at all

## **Household Income**

*Household income is thought to be an important variable in predicting social behaviour. For example, someone with a higher income could be more likely to have donated in the past 12 months.*

- Total household income (INCMHSD, Position 2248)†\*
  - Less than \$40,000
  - \$40,000 to \$79,999
  - \$80,000 or more

## CULTURAL VARIABLES

*Preliminary research showed that each of the 11 cultural activities appeared to have a positive effect on the one or more of the social engagement variables selected.*

- During the past 12 months, as a leisure activity (not for paid work or studies) did you read a newspaper? (LCA\_Q110, Position 1802) †  
→No  
→Yes
- During the past 12 months, as a leisure activity (not for paid work or studies) did you read a book? (LCA\_Q130, Position 1806) †  
→No  
→Yes
- During the past 12 months, did you: go to a movie or drive-in? (LCA\_Q210, Position 1815) †  
→No  
→Yes
- During the past 12 months, did you: attend a concert or performance by professional artists of music, dance, theatre, or opera, excluding cultural festivals? (LCA\_Q300, Position 1833) †  
→No  
→Yes
- During the past 12 months, did you: attend a theatrical performance such as drama, musical theatre, dinner theatre, comedy? (LCA\_Q310, Position 1834) †  
→No  
→Yes
- During the past 12 months, did you: attend a popular musical performance such as pop/rock, jazz, blues, folk, country and western? (LCA\_Q320, Position 1836) †  
→No  
→Yes
- During the past 12 months, did you: attend a symphonic or classical music performance? (LCA\_Q330, Position 1838) †  
→No  
→Yes
- During the past 12 months, did you: go to a cultural or artistic festival (such as film, fringe, dance, jazz, folk, rock, buskers or comedy)? (LCA\_Q340, Position 1840) †  
→No  
→Yes

- During the past 12 months, did you: go to a public art gallery or art museum (including attendance at special art exhibits)? (LCA\_Q410, Position 1846)†  
→No  
→Yes
- During the past 12 months, did you: go to an historic site? (LCA\_Q430, Position 1850)†  
→No  
→Yes
- During the past 12 months, did you: go to a conservation area or nature park? (LCA\_Q450, Position 1854)†  
→No  
→Yes



## Bibliography

*Logistic Regression*, G. David Garson. 2006,  
<http://www2.chass.ncsu.edu/garson/PA765/logistic.htm>

*Logistic Regression*, 2002,  
<http://userwww.sfsu.edu/~efc/classes/biol710/logistic/logisticreg.htm>

*Logistic Regression*, Gerard E. Dallal, Ph.D. 2001, <http://www.tufts.edu/~gdallal/logistic.htm>

*General Social Survey- Cycle 19: Time Use (2005) User's Guide of the Public Use Microdata File*, Statistics Canada, Social and Aboriginal Statistics Division. December 2006,  
<http://data.library.ubc.ca/datalib/survey/statscan/gss/cycle19/gssc19gidv2.pdf>

*The Diversity of Cultural Participation*, Ostrower, Francie. Urban Institute, November 2005,  
<http://www.wallacefoundation.org/NR/rdonlyres/687A0402-4FAC-478B-B6ED-DC96EA0788A8/0/TheDiversityofCulturalParticipation.pdf>