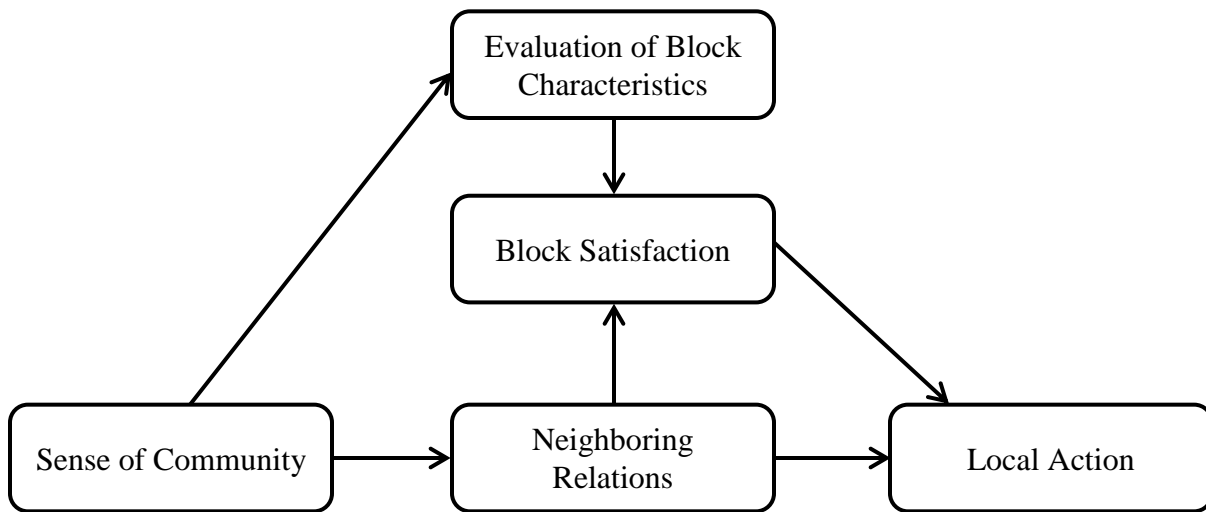
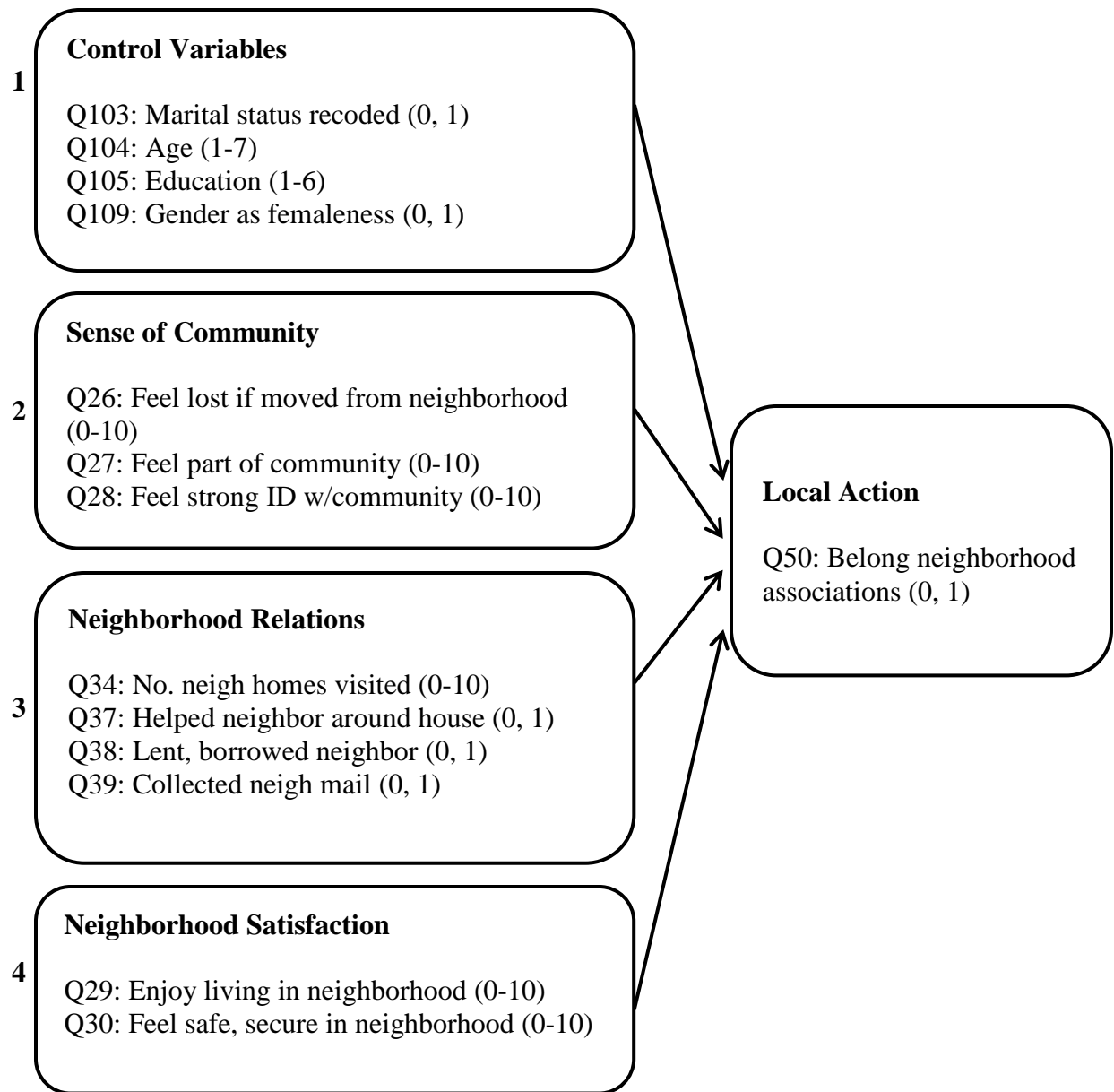


LOGISTIC REGRESSION
Pat Groble & Serineh Baboomian
04/03/13
COM 631/731
National Community Study 2006 data set

1. Model and Theory



General path model for determinants of local action adopted from Chavis and Wandersman, 1990—Chavis, D. M., & Wandersman, A. (1990). Sense of community in the urban environment: a catalyst for participation and community development. *American Journal of Community Psychology*, 18(1), 55-81.



2. Running SPSS

ANALYZE

→ REGRESSION

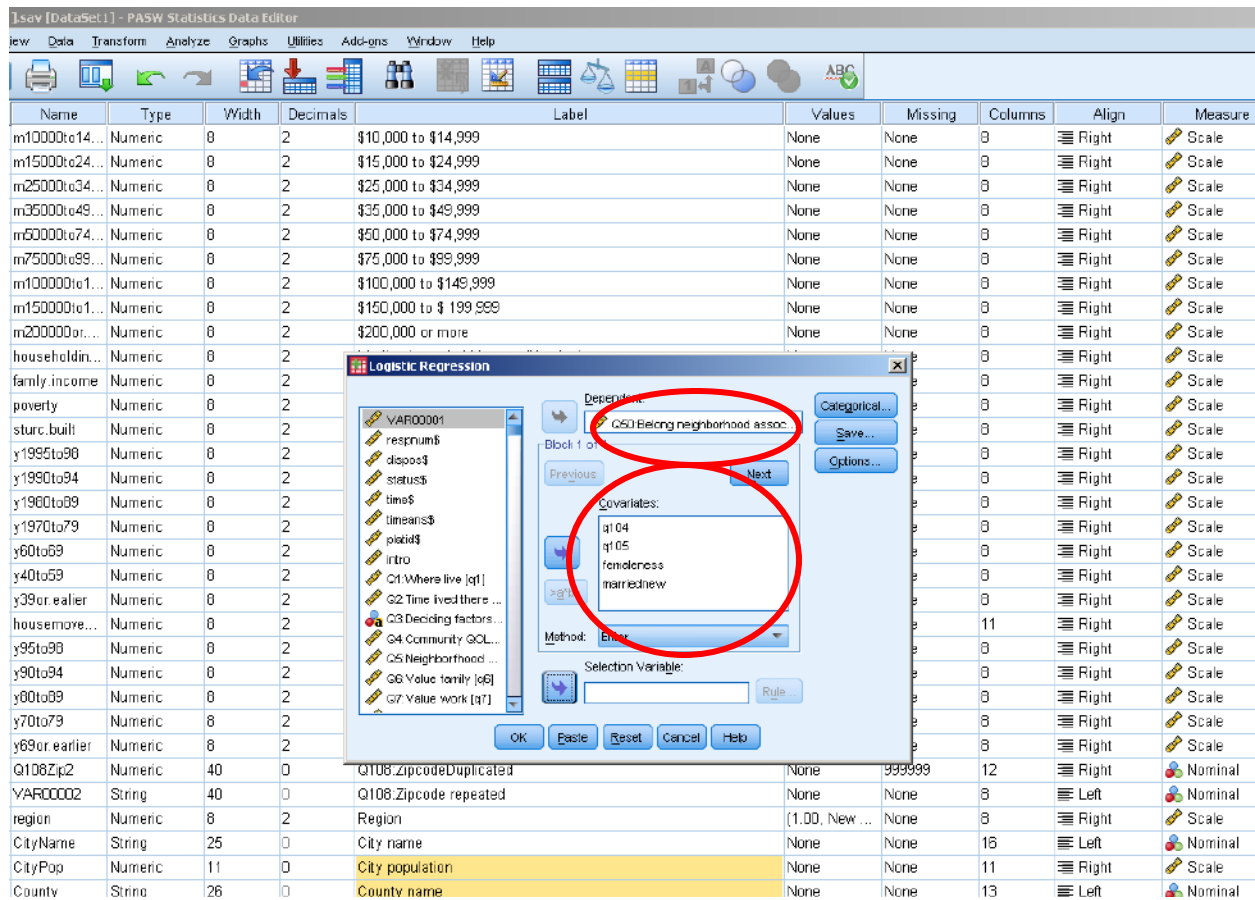
→ BINARY LOGISTIC

DEPENDENT: Insert dependent variable

BLOCK 1

COVARIATES: Insert independent variables for Block 1

METHOD: Enter (Forced entry)



The screenshot shows the SPSS Data Editor window with a list of variables. A dialog box for 'Logistic Regression' is open, showing the following configuration:

- Dependent Variable: G50 Belong neighborhood assoc.
- Block 1 of 1 (Covariates): q104, q105, femaleness, marriednew
- Method: Enter

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
m10000to14...	Numeric	8	2	\$10,000 to \$14,999	None	None	8	Right	Scale
m15000to24...	Numeric	8	2	\$15,000 to \$24,999	None	None	8	Right	Scale
m25000to34...	Numeric	8	2	\$25,000 to \$34,999	None	None	8	Right	Scale
m35000to49...	Numeric	8	2	\$35,000 to \$49,999	None	None	8	Right	Scale
m50000to74...	Numeric	8	2	\$50,000 to \$74,999	None	None	8	Right	Scale
m75000to99...	Numeric	8	2	\$75,000 to \$99,999	None	None	8	Right	Scale
m100000to1...	Numeric	8	2	\$100,000 to \$149,999	None	None	8	Right	Scale
m150000to1...	Numeric	8	2	\$150,000 to \$ 199,999	None	None	8	Right	Scale
m200000or...	Numeric	8	2	\$200,000 or more	None	None	8	Right	Scale
householdin...	Numeric	8	2		None	None	8	Right	Scale
family income	Numeric	8	2		None	None	8	Right	Scale
poverty	Numeric	8	2		None	None	8	Right	Scale
sturc.built	Numeric	8	2		None	None	8	Right	Scale
y1995to98	Numeric	8	2		None	None	8	Right	Scale
y1990to94	Numeric	8	2		None	None	8	Right	Scale
y1960to89	Numeric	8	2		None	None	8	Right	Scale
y1970to79	Numeric	8	2		None	None	8	Right	Scale
y60to69	Numeric	8	2		None	None	8	Right	Scale
y40to59	Numeric	8	2		None	None	8	Right	Scale
y39or earlier	Numeric	8	2		None	None	8	Right	Scale
housemore...	Numeric	8	2		None	None	11	Right	Scale
y95to98	Numeric	8	2		None	None	8	Right	Scale
y90to94	Numeric	8	2		None	None	8	Right	Scale
y80to89	Numeric	8	2		None	None	8	Right	Scale
y70to79	Numeric	8	2		None	None	8	Right	Scale
y69or earlier	Numeric	8	2		None	None	8	Right	Scale
Q108Zipcode	Numeric	40	0	Q108.ZipcodeDuplicated	None	999999	12	Right	Nominal
YAR00002	String	40	0	Q108.Zipcode repeated	None	None	8	Left	Nominal
region	Numeric	8	2	Region	{1.00, New ...	None	8	Right	Scale
CityName	String	25	0	City name	None	None	16	Left	Nominal
CityPop	Numeric	11	0	City population	None	None	11	Right	Scale
County	String	26	0	County name	None	None	13	Left	Nominal

Click NEXT

REPEAT this process for Blocks 2 and 3

BLOCK 4

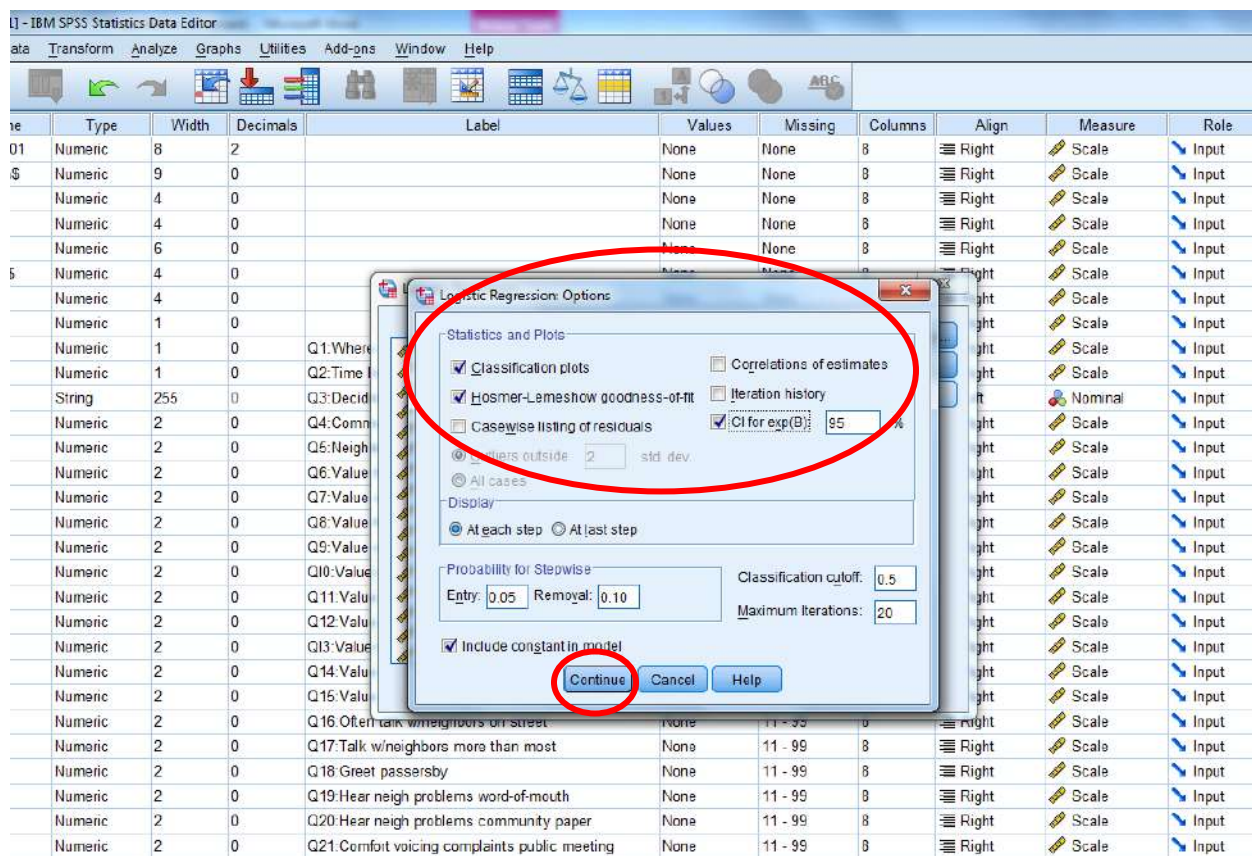
COVARIATES: Insert independent variables for Block 4

Then click on OPTIONS:

Check boxes for Classification Plots

Hosmer-Lemeshow

CI for exp(B) at 95%



Click on CONTINUE

Then click OK

The screenshot shows the SPSS Statistics Data Editor interface. A dialog box titled "Logistic Regression" is open, positioned over a data table. The dialog box has the following settings:

- Dependent:** Q50:Belong neighborhood as...
- Covariates:** q29, q30
- Method:** Enter
- Selection Variable:** (empty)

The "OK" button at the bottom of the dialog box is circled in red. The background data table has columns for Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure, and Role. The data rows are labeled with "eric" and "g".

3. SPSS Output

```
LOGISTIC REGRESSION VARIABLES q50
/METHOD=ENTER q104 q105 femaleness marriednew
/METHOD=ENTER q26 q27 q28
/METHOD=ENTER q34 q37 q38 q39
/METHOD=ENTER q29 q30
/CLASSPLOT
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

Logistic Regression

[DataSet1] C:\Documents and Settings\2532364\Local Settings\Temporary Internet Files\Content.IE5\QHETIL6R\natcom[1].sav

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	411	86.2
	Missing Cases	66	13.8
	Total	477	100.0
Unselected Cases		0	.0
Total		477	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
no	0
yes	1

Block 0: Beginning Block

Classification Table^{a,b}

Observed			Predicted	
			Q50:Belong neighborhood associations	
			no	yes
Step 0	Q50:Belong neighborhood associations	no	327	0
		yes	84	0
Overall Percentage				

a. Constant is included in the model.

b. The cut value is .500

Classification Table^{a,b}

Observed			Predicted	
			Percentage Correct	
Step 0	Q50:Belong neighborhood associations	no	100.0	
		yes	.0	
Overall Percentage			79.6	

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1.359	.122	123.457	1	.000	.257

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	q104	2.395	1	.122
		q105	13.470	1	.000
		femaleness	3.783	1	.052
		marriednew	5.837	1	.016
Overall Statistics			22.687	4	.000

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	25.018	4	.000
	Block	25.018	4	.000
	Model	25.018	4	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	391.254 ^a	.059	.093

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	15.428	8	.051

Contingency Table for Hosmer and Lemeshow Test

		Q50:Belong neighborhood associations = no		Q50:Belong neighborhood associations = yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	37	38.665	4	2.335	41
	2	39	38.659	4	4.341	43
	3	38	36.371	4	5.629	42
	4	35	32.065	3	5.935	38
	5	29	33.639	12	7.361	41
	6	36	33.498	6	8.502	42
	7	35	31.483	6	9.517	41
	8	23	29.672	18	11.328	41
	9	28	26.647	11	12.353	39
	10	27	26.300	16	16.700	43

Classification Table^a

Observed			Predicted	
			Q50:Belong neighborhood associations	
			no	yes
Step 1	Q50:Belong neighborhood associations	no	327	0
		yes	84	0
Overall Percentage				

a. The cut value is .500

Classification Table^a

Observed			Predicted
			Percentage Correct
Step 1	Q50:Belong neighborhood associations	no	100.0
		yes	.0
Overall Percentage			79.6

a. The cut value is .500

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	390.257 ^a	.061	.096

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	10.462	8	.234

Contingency Table for Hosmer and Lemeshow Test

		Q50:Belong neighborhood associations = no		Q50:Belong neighborhood associations = yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	38	38.653	3	2.347	41
	2	35	36.949	6	4.051	41
	3	40	35.688	1	5.312	41
	4	32	34.696	9	6.304	41
	5	33	33.667	8	7.333	41
	6	36	32.711	5	8.289	41
	7	32	31.513	9	9.487	41
	8	30	29.971	11	11.029	41
	9	24	27.972	17	13.028	41
	10	27	25.180	15	16.820	42

Classification Table^a

	Observed		Predicted	
			Q50:Belong neighborhood associations	
			no	yes
Step 1	Q50:Belong neighborhood associations	no	326	1
		yes	84	0
	Overall Percentage			

a. The cut value is .500

Classification Table^a

Observed			Predicted	
			Percentage Correct	
Step 1	Q50:Belong neighborhood associations	no	99.7	
		yes	.0	
Overall Percentage			79.3	

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	q104	-.182	.088	4.316	1	.038	.834
	q105	.339	.108	9.962	1	.002	1.404
	femaleness	-.456	.259	3.102	1	.078	.634
	marriednew	.561	.274	4.189	1	.041	1.752
	q26	-.018	.040	.204	1	.652	.982
	q27	.042	.070	.367	1	.544	1.043
	q28	.010	.067	.021	1	.885	1.010
	Constant	-2.427	.613	15.673	1	.000	.088

a. Variable(s) entered on step 1: q26, q27, q28.

Variables in the Equation

		95% C.I. for EXP(B)	
		Lower	Upper
Step 1 ^a	q104	.702	.990
	q105	1.137	1.734
	femaleness	.381	1.053
	marriednew	1.024	2.997
	q26	.909	1.062
	q27	.910	1.196
	q28	.886	1.151
	Constant		

a. Variable(s) entered on step 1: q26, q27, q28.

Step number: 1

Observed Groups and Predicted Probabilities

y - yes

Each Symbol Represents 2 Cases.

Block 3: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	11.221	4	.024
	Block	11.221	4	.024
	Model	37.237	11	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	379.036 ^a	.087	.136

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.633	8	.955

Contingency Table for Hosmer and Lemeshow Test

		Q50:Belong neighborhood associations = no		Q50:Belong neighborhood associations = yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	39	39.319	2	1.681	41
	2	38	37.566	3	3.434	41
	3	35	36.353	6	4.647	41
	4	35	35.204	6	5.796	41
	5	35	34.150	6	6.850	41
	6	33	32.889	8	8.111	41
	7	34	31.667	7	9.333	41
	8	30	30.023	11	10.977	41
	9	24	27.071	17	13.929	41
	10	24	22.757	18	19.243	42

Classification Table^a

Observed			Predicted	
			Q50:Belong neighborhood associations	
			no	yes
Step 1	Q50:Belong neighborhood associations	no	323	4
		yes	81	3
Overall Percentage				

a. The cut value is .500

Classification Table^a

Observed			Predicted
			Percentage Correct
Step 1	Q50:Belong neighborhood associations	no	98.8
		yes	3.6
Overall Percentage			79.3

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	q104	-.188	.090	4.344	1	.037	.828
	q105	.287	.109	6.959	1	.008	1.333
	femaleness	-.372	.267	1.940	1	.164	.690
	marriednew	.478	.280	2.916	1	.088	1.612
	q26	-.020	.040	.253	1	.615	.980
	q27	-.011	.073	.023	1	.879	.989
	q28	-.005	.069	.005	1	.943	.995
	q34	.106	.045	5.509	1	.019	1.111
	q37	.370	.438	.715	1	.398	1.448
	q38	.028	.464	.004	1	.952	1.029
	q39	.346	.302	1.314	1	.252	1.413
	Constant	-2.767	.723	14.646	1	.000	.063

a. Variable(s) entered on step 1: q34, q37, q38, q39.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	374.817 ^a	.096	.151

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	11.134	8	.194

Contingency Table for Hosmer and Lemeshow Test

		Q50:Belong neighborhood associations = no		Q50:Belong neighborhood associations = yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	39	39.518	2	1.482	41
	2	38	37.951	3	3.039	41
	3	36	36.761	5	4.239	41
	4	34	35.361	7	5.639	41
	5	38	34.199	3	6.801	41
	6	33	32.998	8	8.002	41
	7	36	31.384	5	9.616	41
	8	27	29.597	14	11.403	41
	9	21	26.740	20	14.260	41
	10	25	22.479	17	19.521	42

Classification Table^a

	Observed		Predicted	
			Q50:Belong neighborhood associations	
			no	yes
Step 1	Q50:Belong neighborhood associations	no	319	8
		yes	81	3
	Overall Percentage			

a. The cut value is .500

Classification Table^a

Observed			Predicted
			Percentage Correct
Step 1	Q50:Belong neighborhood associations	no	97.6
		yes	3.6
Overall Percentage			78.3

a. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a						
q104	-.154	.092	2.797	1	.094	.857
q105	.331	.112	8.719	1	.003	1.393
femaleness	-.332	.270	1.515	1	.218	.717
marriednew	.495	.282	3.094	1	.079	1.641
q26	-.014	.042	.116	1	.733	.986
q27	.019	.080	.058	1	.809	1.020
q28	.017	.073	.057	1	.812	1.018
q34	.113	.046	5.979	1	.014	1.119
q37	.327	.445	.540	1	.462	1.387
q38	.018	.474	.001	1	.969	1.018
q39	.397	.303	1.717	1	.190	1.488
q29	-.078	.087	.816	1	.366	.925
q30	-.089	.069	1.665	1	.197	.914
Constant	-2.156	.764	7.969	1	.005	.116

a. Variable(s) entered on step 1: q29, q30.

4. Tables of Results

Table 1: Prediction of Belonging to Neighborhood Organizations via Logistic Regression.

				Model Statistics					
	r	Exp (B) in	Final Exp (B)	Block Chi-Sq	Model Chi-Sq	Model - 2LL	Cox & Snell R ²	Nag. R ²	Hosmer & Lemeshow Chi-Sq
Block 1: Control				25.018**	25.018**	391.254	.059	.093	10.462
marriednew	.112*	1.783*	1.641						
Q104	-.089	.848*	.857						
Q105	.141**	1.421**	1.393**						
femaleness	-.080	.631	.717						
Block 2: Sense of Comm.				.997	26.016**	390.257	.061	.096	4.940
Q26	-.068	0.982	.986						
Q27	.006	1.043	1.020						
Q28	.009	1.010	1.018						
Block 3: Soc. Rels.				11.221*	37.237**	379.036	.087	.136	2.633
Q34	.163**	1.111*	1.119*						
Q37	.110*	1.448	1.387						
Q38	.109*	1.029	1.018						
Q39	.147**	1.413	1.488						
Block 4: Neighbor. Satisfaction				4.219	41.456**	374.817	.096	.151	11.134
Q29	-.068	.925	.925						
Q30	-.080	.914	.914						

* p < .05; **p < .01

Table 2: Classification Results

		Final Predicted Group (Blocks 1 – 4)			
		Q50			Percentage Correct
		No	Yes		
Block 4	Q50 Belong to neighborhood associations	No	319	8	97.6
		Yes	81	3	3.6
Overall percentage					78.3

Press' Q Calculation:

Formula: $[N-(nK)]^2 / N(K-1)$

Where N= total sample size

n=number of observations correctly classified

K=number of groups

In our model:

$$N = 411$$

$$n = 322 \quad [319+3]$$

$$K = 2$$

$$[411-(322*2)]^2 / 411(2-1)$$

$$54289/411$$

$$\text{Press' Q} = 132.09$$

$$df = 1$$

$$\text{Chi-sq}_{\text{crit}} = 10.83; \quad p = .001$$

5. Findings

We used logistic regression to predict the likelihood of someone belonging to a neighborhood association, based on variables dealing with social relations, a sense of community and a sense of belonging.

Table 1 summarizes the findings from the logistic regression. Only two blocks had significant Chi-squares: Block 1 (Control) and Block 3 (Social Relations). The total equation with four blocks is significant at $p < .01$. Each Exp(B) indicates a decrease or increase in the odds of the occurrence of the dependent value, assuming all other IVs are controlled for. Within

the two significant blocks, only two independent variables have significant unique contributions to the prediction of belonging to neighborhood associations. Both of these Exp(B)s are positive, indicating that each increases the odds of the dependent variable occurring (here, belonging to a neighborhood association), when controlling for all other independent variables in the model. In terms of education, with a Final Exp(B) of 1.393, this means that for every increase in one unit of education, there will be a predicted increase of 39% in the odds of belonging to a neighborhood association occurring, when all other independent variables are controlled. The other significant predictor variable, number of close neighborhood homes visited, had a Final Exp(B) of 1.119, indicating that for every additional house visited, the odds of belonging to a neighborhood association occurring increases by 11.9%, when all other independent variables are controlled. All our Hosmer and Lemeshow Chi-squares were non-significant, which is what we want.

Table 2 shows that our model correctly predicted 78.3% of the classification. This is supported by the Press' Q and Critical chi-square statistics. Press' Q was calculated at 132.09. Critical Chi-square equals 10.83. Since our Press' Q exceeds critical chi-square, the accuracy of our predictions is greater than what could be expected by chance