**Appendix A. Frequency Tables**

**Table A1. Race and Ethnicity of the Respondents in IIMMLA Dataset**

|  |  |  |
| --- | --- | --- |
|  | Frequency | Percent |
| Mexican | 1244 | 26.7 |
| Salvadoran/Guatemalan | 376 | 8.1 |
| Other Latino | 188 | 4 |
| Chinese | 400 | 8.6 |
| Korean | 401 | 8.6 |
| Vietnamese | 401 | 8.6 |
| Filipino | 401 | 8.6 |
| Other Asian | 95 | 2 |
| White | 704 | 15.1 |
| Black | 445 | 9.6 |
| Total | 4655 | 100 |

**Table A2. LCA Variables by Race and Ethnicity: Voter Registration**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Registered | Somewhere else | Not registered | Total |
| Mexican | 785 | 69 | 218 | 1072 |
| Salvadoran/  Guatemalan | 222 | 13 | 45 | 280 |
| Other Latino | 128 | 12 | 16 | 156 |
| Chinese | 257 | 29 | 93 | 379 |
| Korean | 232 | 29 | 81 | 342 |
| Vietnamese | 259 | 22 | 90 | 371 |
| Filipino | 251 | 27 | 81 | 359 |
| Other Asian | 47 | 7 | 25 | 79 |
| White | 523 | 55 | 97 | 675 |
| Black | 336 | 40 | 61 | 437 |
| Total | 3040 | 303 | 807 | 4150 |

**Table A3. LCA Variables by Race and Ethnicity: Protest Participation**

|  |  |  |  |
| --- | --- | --- | --- |
|  | No | Yes | Total |
| Mexican | 1056 | 186 | 1242 |
| Salvadoran/  Guatemalan | 322 | 53 | 375 |
| Other Latino | 159 | 29 | 188 |
| Chinese | 358 | 42 | 400 |
| Korean | 361 | 40 | 401 |
| Vietnamese | 373 | 28 | 401 |
| Filipino | 375 | 24 | 399 |
| Other Asian | 85 | 10 | 95 |
| White | 598 | 106 | 704 |
| Black | 381 | 63 | 444 |
| Total | 4068 | 581 | 4649 |

**Table A4. LCA Variables by Race and Ethnicity: Understanding Politics**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Very Low | Low | High | Very High | Total |
| Mexican | 51 | 112 | 541 | 533 | 1237 |
| Salvadoran/  Guatemalan | 10 | 31 | 144 | 189 | 374 |
| Other Latino | 4 | 14 | 69 | 100 | 187 |
| Chinese | 11 | 53 | 195 | 139 | 398 |
| Korean | 6 | 40 | 184 | 171 | 401 |
| Vietnamese | 18 | 52 | 193 | 135 | 398 |
| Filipino | 10 | 39 | 197 | 153 | 399 |
| Other Asian | 3 | 7 | 46 | 39 | 95 |
| White | 18 | 45 | 294 | 341 | 698 |
| Black | 19 | 38 | 168 | 217 | 442 |
| Total | 150 | 431 | 2031 | 2017 | 4629 |

**Table A5. LCA Variables by Race and Ethnicity: Experience of Discrimination**

|  |  |  |  |
| --- | --- | --- | --- |
|  | No | Yes | Total |
| Mexican | 841 | 400 | 1241 |
| Salvadoran/  Guatemalan | 248 | 126 | 374 |
| Other Latino | 134 | 54 | 188 |
| Chinese | 280 | 117 | 397 |
| Korean | 243 | 157 | 400 |
| Vietnamese | 248 | 151 | 399 |
| Filipino | 268 | 132 | 400 |
| Other Asian | 57 | 37 | 94 |
| White | 544 | 157 | 701 |
| Black | 194 | 249 | 443 |
| Total | 3057 | 1580 | 4637 |

**Table A6. LCA Variables by Race and Ethnicity: Education**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | No HSD | HSD | Some College | BA | Grad | Total |
| Mexican | 286 | 334 | 428 | 137 | 59 | 1244 |
| Salvadoran/Guatemalan | 59 | 80 | 175 | 46 | 16 | 376 |
| Other Latino | 14 | 27 | 93 | 36 | 18 | 188 |
| Chinese | 2 | 21 | 119 | 172 | 86 | 400 |
| Korean | 9 | 25 | 127 | 168 | 72 | 401 |
| Vietnamese | 8 | 30 | 172 | 149 | 42 | 401 |
| Filipino | 12 | 35 | 185 | 129 | 40 | 401 |
| Other Asian | 10 | 10 | 32 | 33 | 10 | 95 |
| White | 54 | 107 | 239 | 196 | 108 | 704 |
| Black | 67 | 86 | 202 | 66 | 24 | 445 |
| Total | 521 | 755 | 1772 | 1132 | 475 | 4655 |

**Appendix B. Logistic Regression Models**

I use logistic regression models to demonstrate how the principle of “different models for different groups” functions in practice. By running separate regressions for each racial and ethnic group, we do discover that different groups have different political participation models, yet we do not understand how those models are related to each other. The two indicators of political participation—voter registration status and experience of protesting in the streets—are used as dependent variables, while three other variables—education, “understanding politics,” and experience of racial discrimination—are included in the model as independent variables.[[1]](#footnote-1) The models can be stated as:

Whereas P stands for the probability that a given respondent is registered as a voter (or having participated in a street protest), , , and , respectively, represent education level, understanding of politics, and experience of discrimination, all measured in the response categories presented in Table 1 of the article. represents error terms. The two models are estimated for the whole data set and then separately for ten racial and ethnic groups, resulting in twenty-two separate logistic regression models in total.

**Table B1. The Results from the Logistic Regression Models for Voter Registration**

|  |  |  |  |
| --- | --- | --- | --- |
| Voter Registration | Education | Understanding Politics | Experience of Discrimination |
|  |  |  |  |
| All | 0.29†  (0.03) | 0.48†  (0.05) | 0  (0.09) |
|  |  |  |  |
| Mexican | 0.59†  (0.07) | 0.06  (0.10) | -0.16  (0.17) |
| Salvadoran/Guatemalan | 0.16  (0.14) | 0.64†  (0.20) | -0.21  (.35) |
| Other Latino | 0.70†  (0.24) | 0.70\*  (0.33) | -1.06  (0.60) |
| Chinese | 0.18\*  (0.10) | 0.50†  (0.15) | -0.32  (0.27) |
| Korean | 0.53†  (0.10) | 0.72†  (0.19) | 0.08  (0.28) |
| Vietnamese | 0.17  (0.11) | 0.70†  (0.16) | 0.63\*  (0.28) |
| Filipino | 0.35†  (0.11) | 0.69†  (0.18) | 0.05  (0.29) |
| Other Asian | 0.10  (0.18) | 0.06  (0.38) | 0.19  (0.51) |
| Black | 0.52†  (0.13) | 0.34\*  (0.16) | 0.19  (0.29) |
| White | 0.60†  (0.09) | 0.57†  (0.15) | 0.07  (0.28) |

\*Standard errors are presented within parentheses. Significance levels are indicated as:

†p<0.01, \*p<0.05.

**Table B2. The Results from the Logistic Regression Models for Protest Participation**

|  |  |  |  |
| --- | --- | --- | --- |
| Protest | Education | Understanding Politics | Experience of Discrimination |
|  |  |  |  |
| All | 0.12†  (0.03) | 0.42†  (0.07) | 0.65†  (0.09) |
|  |  |  |  |
| Mexican | 0.18†  (0.06) | 0.28\*  (0.12) | 0.76†  (0.16) |
| Salvadoran/Guatemalan | 0.38†  (0.11) | 0.41  (0.25) | 0.69\*  (0.31) |
| Other Latino | 0.18  (0.15) | 0.29  (0.31) | -0.26  (0.47) |
| Chinese | 0.12  (0.14) | 0.72†  (0.27) | 0.91†  (0.33) |
| Korean | 0.15  (0.14) | 0.22  (0.25) | 0.82\*  (0.34) |
| Vietnamese | 0.06  (0.17) | 0.41  (0.28) | 0.79\*  (0.40) |
| Filipino | 0.03  (0.17) | -0.01  (.30) | 0.43  (0.43) |
| Other Asian | 0.24  (0.24) | 0.80  (0.61) | 1.36  (0.74) |
| Black | 0.26†  (0.10) | 0.29  (0.20) | 1.23†  (0.33) |
| White | 0.25†  (0.07) | 0.53†  (0.18) | 0.35  (0.25) |

\*Standard errors are presented within parentheses. Significance levels are indicated as:

†p<0.01, \*p<0.05.

Tables B1 and B2 summarize the results from the two sets of logistic regression models. The “all” columns in both Tables B1 and B2 indicate that when applied to the entire data set, regardless of race and ethnicity, the logistic regression models confirm the findings from previous research: educated respondents with political knowledge are more likely to be registered; and educated, knowledgeable respondents who experienced racial discrimination tend to participate in street protest more often. These results are far from surprising.

When we break the results down by racial and ethnic groups, however, we see “different models for different groups.” The models vary widely by racial and ethnic group, and we cannot discern why they differ from each other in unexpected ways. For example, in Table B1, many groups follow the pattern shown in the “all” model, with education and “understanding politics” having positive, statistically significant coefficients on the odds of being registered. As we have seen in the literature on minority politics, however, there are exceptions: education is not significant for Salvadoran/Guatemalan, Vietnamese, and “Other Asian” groups; understanding of politics does not matter for Mexican and “Other Asian” groups; and experience of racial discrimination matters only for Vietnamese.

The results are even more confusing in the models for street protest, presented in Table B2. Education matters for Mexican, Salvadoran/Guatemalan, Black, and White, but not for other groups; and understanding of politics matters for Mexican, Chinese, and White, but not for others. Experience of discrimination generally increases the odds of having participated in street protest, but not for “other Latino,” Filipino, “Other Asian,” and White. In short, as we have seen in the review of the literature, the findings from the standard SES model are generally confirmed, but with many exceptions, and we do not know why some groups fit into the model and others do not.

**Appendix C. Regression Analysis and Latent Class Model**

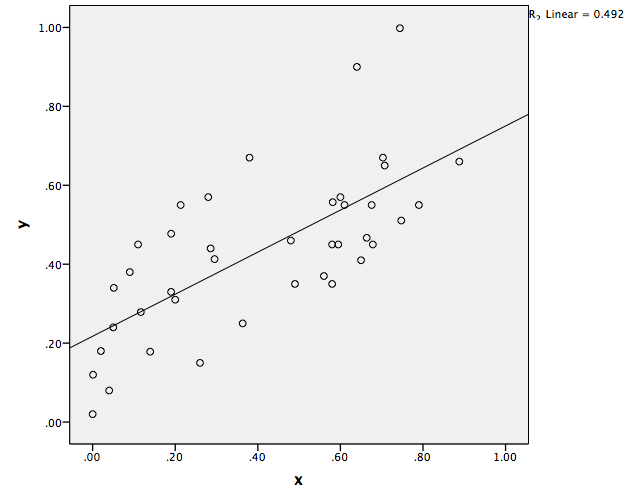
Figure C1. A Hypothetical Data Set with a Regression Line

Figure C1 presents a scatterplot of data points from a randomly generated hypothetical data set, consisting of two continuous variables (x, y). Suppose that this hypothetical data set is a simplified version of our data set, and x and y each represents education and political participation, respectively.[[2]](#footnote-2)

In order to examine the relationship between the two variables, we would regress y to x, using x as an independent variable and y as the dependent variable. The analysis will produce a regression line, as presented in the figure.

The line fits the data relatively well (R2=0.492), as suggested by the general outlook of the dots in the scatterplot. Although there are some exceptions (i.e. two data points in the upper-right side), we can safely state that higher education level correlates with higher likelihood of participation.

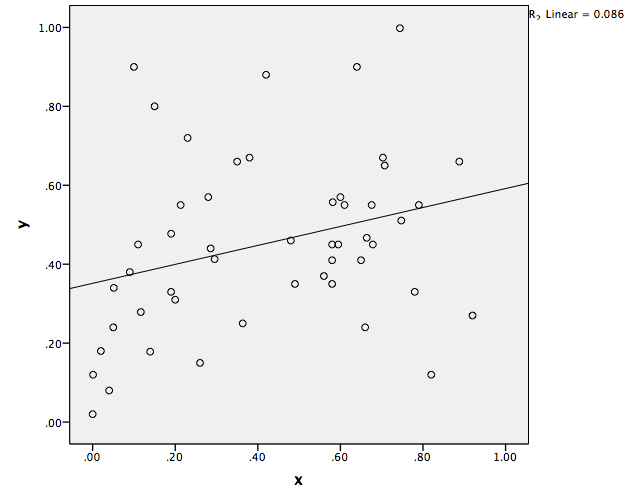
Figure C2. A Hypothetical Data Set with a Regression Line (2)

Figure C2 represents the result of the same exercise, this time using a different hypothetical data set, also randomly generated. In this case, as the scatterplot suggests, data points are spread out more widely across the plot, and the regression line fits less well into the overall distribution (R2=0.086). There are considerable number of data points far away from the regression line, especially at the top and bottom of the scatterplot.

We could still state that higher education level leads to higher likelihood of political participation, given the slope of the regression line, yet there is a considerable number of exceptions to the general statement. For instance, the data points in the upper-left corner, which represent the respondents with low education level and high likelihood of political participation, and those in the bottom-right corner, who represent those with high education level and low likelihood of participation, are such exceptions. We could perhaps state that this model fits the data, but not as well as the previous model in Figure C1, and clearly with some exceptions.

In the literature review, we have seen many works on minority politics that follow this pattern: the model seems to fit, but not as well as the standard SES model derived from an all-White sample, and clearly with many exceptions.

Figure C3. Latent Class Model

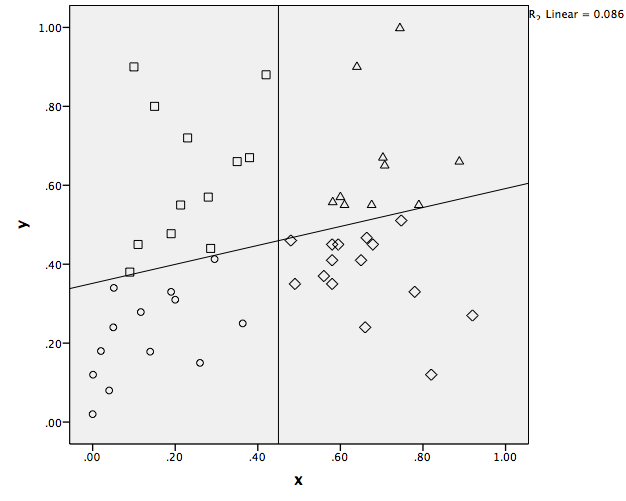


Figure C3 represents how the results from LCA would map onto the scatterplot presented in Figure C2. The LCA algorithm divides the respondents into four categories, or four latent classes. Circles (low education/low participation) and triangles (high education and high participation) can easily be derived from the regression models presented above; squares (low education/high participation) and rhombuses (high education/low participation), however, are new additions. These respondents were regarded as exceptions in the previous models, and effectively neglected in the general summary of the finding (i.e. education increases the likelihood of participation).

In sum, with Latent Class Models, we can effectively detect the exceptions neglected in regression models and capture the heterogeneity within a data set. To be clear, LCA cannot substitute regression models entirely; rather, the two techniques provide different outlooks on the same data, each with a different focus. In this article, I choose LCA to capture intra-group differences and inter-group commonalities.

**Appendix D. Estimates of Fit from Latent Class Models**

Table D1. Estimates of Fit from Latent Class Models

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | LL | BIC (LL) | Npar | Chi-square | df | p-value |
| Model1 | 1 Class | -4610.33 | 9245.618 | 3 | 622.7216 | 237 | 2.00E-36 |
| Model2 | 2 Class | -4439.95 | 9013.022 | 16 | 281.9702 | 224 | 0.0052 |
| **Model3** | **3 Class** | **-4406.76** | **9054.783** | **29** | **215.574** | **211** | **0.4** |
| **Model4** | **4 Class** | **-4393.53** | **9136.493** | **42** | **189.1282** | **198** | **0.66** |
| Model5 | 5 Class | -4385.68 | 9228.94 | 55 | 173.4187 | 185 | 0.72 |

Table D1 shows the estimates of fit from latent class models, ranging from 1 to 5 classes. The class number denotes the optimal number of types of configuration for the variables described in Table 1 of the article.

In choosing the model, I evaluate several criteria including log likelihood (LL), Bayesian Information Criteria (BIC), and chi-square. Although BIC points toward Model 2 as the most parsimonious model with the least possible number of parameters, I choose Model 4 both in terms of chi-square probability and its substantial results. That is, I value explanatory power of the model over parsimoniousness, given the complex nature of the data set that includes 4655 respondents within ten nested categories of groups (for specific criteria on selecting a model in LCA, see Vermunt and Magdison, 2002).

I choose to focus on Model 4 based on chi-square estimation: when moving from Model 3 to 4, the goodness-of-fit chi-square decreases by approximately 26 (from 215.574 to 189.1282) in exchange for 13 degrees of freedom (from 211 to 198). The difference between the two models is highly significant (p=0.0147<0.02), whereas the difference between Model 4 and 5 is not as clear (p<0.2). In all of the models the data from 4108 out of 4655 total respondents were analyzed due to missing data.

**Appendix E: The Results from 3-class and 5-class LCA models.**

Table E1. Profile of the 3-class Model

|  |  |  |  |
| --- | --- | --- | --- |
|  | Class 1 | Class 2 | Class 3 |
| Class Size | 0.43 | 0.29 | 0.28 |
|  |  |  |  |
| ***Indicators*** |  |  |  |
| **Voter Registration** |  |  |  |
| Yes | 0.53 | 0.94 | 0.84 |
| Somewhere | 0.06 | 0.06 | 0.11 |
| No | 0.41 | 0.00 | 0.05 |
| **Protest** |  |  |  |
| Yes | 0.05 | 0.11 | 0.27 |
| No | 0.95 | 0.89 | 0.73 |
|  |  |  |  |
| ***Covariates*** |  |  |  |
| **Understanding Politics** |  |  |  |
| Strongly Agree | 0.27 | 0.52 | 0.64 |
| Agree | 0.51 | 0.45 | 0.31 |
| Disagree | 0.15 | 0.03 | 0.05 |
| Strongly Disagree | 0.06 | 0.00 | 0.01 |
| **Experience of Discrimination** |  |  |  |
| Yes | 0.32 | 0.02 | 0.72 |
| No | 0.68 | 0.98 | 0.28 |
| **Education** |  |  |  |
| No HSD | 0.19 | 0.03 | 0.01 |
| HSD | 0.22 | 0.10 | 0.12 |
| 1-2yrs of college | 0.28 | 0.32 | 0.29 |
| 3-4yrs of college | 0.10 | 0.07 | 0.09 |
| BA | 0.18 | 0.33 | 0.31 |
| Grad | 0.03 | 0.15 | 0.19 |

The 3-class Model further breaks down registered voters along their means of political participation. In Class 3, we see a class for those who participate in protest as well as elections. This group is highly educated and has the highest level of confidence in their understanding of political issues. In addition, they are most likely to have experienced discrimination based on their racial or ethnic identity (0.72 for “yes” in “experience of discrimination”).

Table E2. Profile of the 5-class Model

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 |
| Class Size | 0.43 | 0.17 | 0.17 | 0.14 | 0.09 |
|  |  |  |  |  |  |
| ***Indicators*** |  |  |  |  |  |
| **Voter Registration** |  |  |  |  |  |
| Yes | 0.91 | 0.55 | 0.64 | 0.47 | 0.82 |
| Somewhere | 0.06 | 0.00 | 0.15 | 0.07 | 0.16 |
| No | 0.04 | 0.44 | 0.21 | 0.46 | 0.03 |
| **Protest** |  |  |  |  |  |
| Yes | 0.14 | 0.09 | 0.02 | 0.02 | 0.53 |
| No | 0.86 | 0.91 | 0.98 | 0.98 | 0.47 |
|  |  |  |  |  |  |
| ***Covariates*** |  |  |  |  |  |
| **Understanding Politics** |  |  |  |  |  |
| Strongly Agree | 0.65 | 0.25 | 0.12 | 0.35 | 0.59 |
| Agree | 0.35 | 0.45 | 0.65 | 0.49 | 0.32 |
| Disagree | 0.00 | 0.25 | 0.23 | 0.01 | 0.08 |
| Strongly Disagree | 0.00 | 0.05 | 0.00 | 0.14 | 0.01 |
| **Experience of Discrimination** |  |  |  |  |  |
| Yes | 0.21 | 0.56 | 0.49 | 0.00 | 0.83 |
| No | 0.79 | 0.44 | 0.51 | 1.00 | 0.17 |
| **Education** |  |  |  |  |  |
| No HSD | 0.03 | 0.24 | 0.05 | 0.23 | 0.00 |
| HSD | 0.14 | 0.17 | 0.11 | 0.32 | 0.06 |
| 1-2yrs of college | 0.28 | 0.42 | 0.24 | 0.17 | 0.43 |
| 3-4yrs of college | 0.09 | 0.09 | 0.06 | 0.12 | 0.07 |
| BA | 0.31 | 0.08 | 0.39 | 0.15 | 0.27 |
| Grad | 0.16 | 0.00 | 0.15 | 0.00 | 0.17 |

The 5-class Model, on the other hand, yields very complicated results. There are three groups coming out of “inactive participants” who are less likely to be registered (Classes 2, 3, and 4) than other classes. However, it is hard to make sense of how five classes compare to each other in other characteristics, given complicated profiles in education and understanding of politics. In summary, the 4-class Model presented in the article reveals not only the most fitting picture, but also the most legible heuristic in terms of understanding the variations in the five variables.

1. The voter registration has been recoded as “yes” (“registered where lives” and “registered somewhere else”) and “no” (“not registered”). I intentionally keep the models simple to underscore the point that the principle of “different models for different groups” provides only limited insight on heterogeneity within groups. If simple, four-variable regression models can generate this much variation among groups, adding more variables will likely make the groups more different (i.e. in terms of model specifications), without providing hints on how to contextualize the differences. [↑](#footnote-ref-1)
2. For the purpose of illustration, x and y are supposed as continuous variables, ranging from 0 to 1. [↑](#footnote-ref-2)