

Kim Davis: Because Jesus Says So

A Look At Media Bias In The United States

Durham, Nichols, Williams, Soto – Oct 4 2015

A Quick Introduction

Rowan county clerk in Kentucky

Denied same-sex couples marriage licenses

Had a secret meeting with the pope



Our Hypothesis

Is media bias uneven when comparing southern news outlets to other news outlets in the United States?

Null: $\text{Bias}(\text{other}) = \text{Bias}(\text{south})$

Alternative: $\text{Bias}(\text{other}) \neq \text{Bias}(\text{south})$

Variables

X	Explanatory
REGION (CATEGORICAL)	South or Other (West, Northeast)
REACH (CATEGORICAL)	Local or Regional
count	Word count of the article
Y	Response
Bias	Binary (0 no event or 1 event)

Other Questions

Region important when considering reach, length of article and gender?

Is there a valid interaction between the explanatory variables of word count, gender or reach?

Study Considerations

Population: Regional and local news outlets in the United States, specifically in California, Washington, New York, Oregon, Nevada, Alabama, Texas, Kentucky, Louisiana

Subjects: individual **opinion based** articles from male and female writers

Excluded all neutral and polarizing (GayStar News, Christian Monitor, Associated Press)

Sampling Plan

Stratification

60 observations gathered: 30 articles in South region, 30 articles in other region

Other region: California, Oregon, Washington, Nevada

Segmented by reach: 15 local articles, 15 national articles

South region: Alabama, Kentucky, Texas, Louisiana

Segmented by reach: 15 local articles, 15 national articles

40 total selected

Sampling Plan

Randomly select 10 (out of 15) articles from each reach strata

Randomize selection using a simple numpy `random.sample()` without replacement

The Analysis

- 1) Does headline predisposition reader?
- 2) Is the article unbalanced from an evidence standpoint?
- 3) Does the article directly slander or compliment the subject or important interests associated with subject matter? (e.g. one example was LGBT activists were called 'militants')
- 4) Are the visuals overly positive or negative?
- 5) Overuse of buzzwords and categorizations? E.g. LGBT, Christian Agenda, Homosexual Agenda

Original Plan: each member conducts five point analysis of each article, grades article based on binary value 0 (unbiased) or 1 (biased)

Consensus rules for article bias indicator (3/4 member grade 1 then 1)

The Analysis – Regional Effect

P(Bias) = REGION

Due to a binary response value, we chose to use binary logistic regression for analysis

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Location	1	4.7626	0.0291

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.3466	0.3446	1.0115	0.3145
Location	Other	1	0.7520	0.3446	4.7626	0.0291

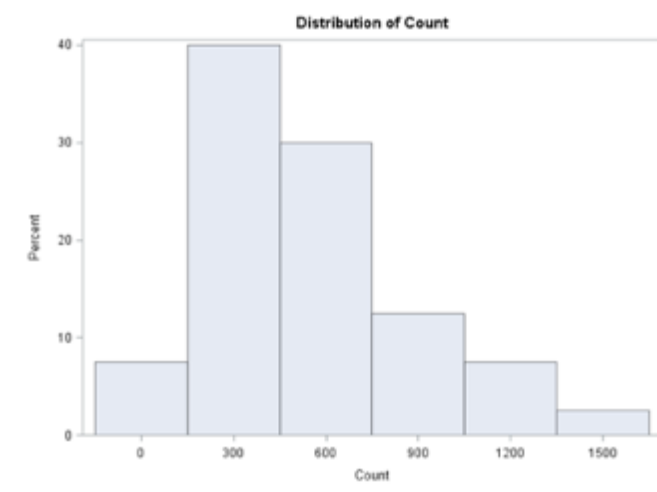
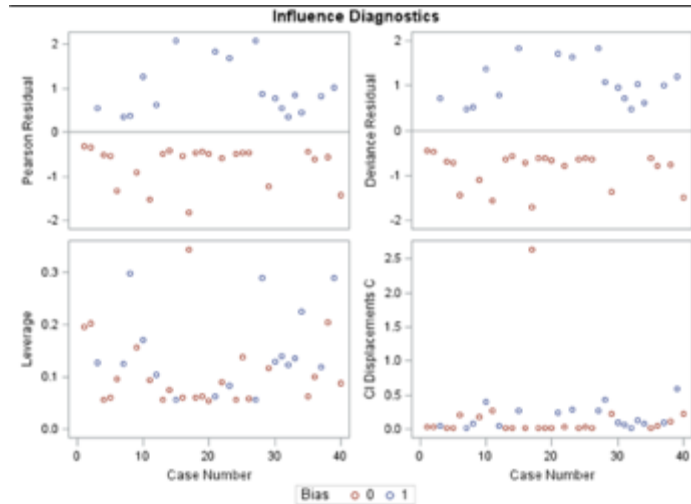
Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	56.548	53.414
SC	58.237	56.792
-2 Log L	54.548	49.414

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Location Other vs South	4.500	1.166	17.373

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	46.0	Somers' D	0.358
Percent Discordant	10.2	Gamma	0.636
Percent Tied	43.7	Tau-a	0.179

The Analysis – With More Cowbell

$$P(\text{Bias}) = \text{count} + \text{REGION} + \text{REACH} + (\text{count} * \text{REACH})$$



The Analysis – With More Cowbell

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	0.8765	0.9607	0.8324	0.3616
Location	Other	1	0.8933	0.4269	4.3794	0.0364
Reach	Local	1	-1.8426	0.9202	4.0091	0.0453
Count		1	-0.00172	0.00145	1.4011	0.2365
Count*Reach	Local	1	0.00298	0.00134	4.9827	0.0256

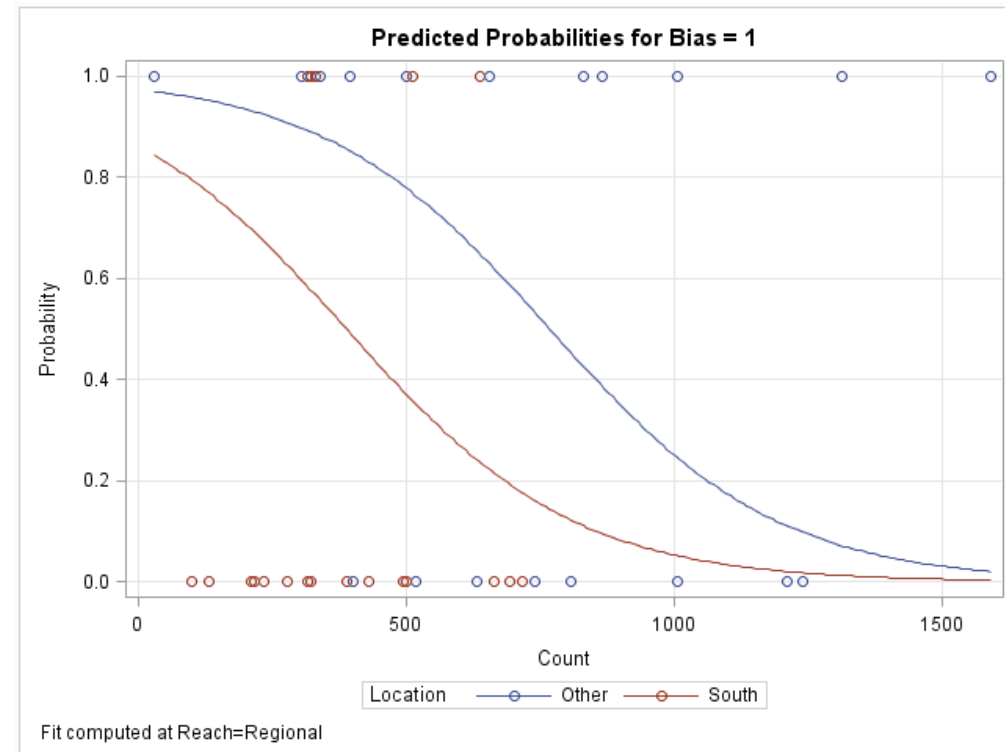
Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Location Other vs South	5.970	1.120	31.819

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
10.6213	8	0.2241

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	56.548	52.530
SC	58.237	60.975
-2 Log L	54.548	42.530

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	79.8	Somers' D	0.609
Percent Discordant	18.9	Gamma	0.617
Percent Tied	1.3	Tau-a	0.305
Pairs	391	c	0.804

The Analysis – Effect Plot of Count By Region



Future Plans

AlchemyAPI - Sentiment Analysis (still owe Raghu a beverage)

Add gender, writer history variables based on cafe discussions

Change to MLR