

Analaysis of Public Housing in County Census Tracts and Proximity to CE(s)

Original Analysis

By: D. White

Date: 2019-11-13

KNITR File: Report_Logistic_Model.Rnw

```
## [1] "Generated on: 2020-04-15 10:12:14"
```

Methods: A multistep data processing routine was performed for each county to integrate spatial and tabular data sets of CE presence/absence and public housing counts at a census tract level in each county. Data were aggregated to the final observed year for each county. Census tract data and HUD data date to 2000. HUD data include Section-8 vouchers and Low-Income Housing Tax Credit (LIHTC) unit counts in each census tract. The final tabular data set for each county is presented in this analysis. A logistic regression was performed on an all county data set to test the probability of a CE located in census tract dependent upon HUD housing counts (Tables 1-4). Tables 5-34 show results of the logistic regression for each individual county. ALB and DGL were not included in the analysis due small sample size. Results for LEB is of particular interest.

The data processing by county can be found in the path below.
C:\Users\whitedl\Documents\R_Code\HUD_Project_Code\CNTYNAME_Public_Housing_CE.R

Table 1 shows the coefficient estimates and related information that result from fitting a logistic regression to predict the probability of a CE located in a census tract based on the predictors Section 8 and Tax Credit housing counts in the corresponding census tracts. A positive coefficient will indicate that an increase in public housing is associated with an increase in the probability of a CE in a given Census Tract. A negative coefficient will imply that increased numbers of public housing units are associated with a decreased probability of a CE in a census tract. A highly significant negative effect is observed for Section 8 housing ($P < .01$) while a positive significant effect ($P < .05$) is observed for Tax Credit housing.

Table 2 reports a NULL model deviance of 1093. A residual deviance of 1032 is observed.

A McFadden statistic is reported in table 3. The McFadden statistic is complementary to a linear regression R^2 value. A value of 0.068 indicates a relatively poor model fit. Although not as severe if it were an OLS R^2 value.

Observations: 1. Many zeros are present in the Tax Credit housing. Originally, these were coded as NA, but that was an error. 2. I examined other model derivatives such as each predictor alone and an interaction effect. Those other models were even less compelling. 3. Overall, I do not find this to be a strong model. But my limited experience with logistic/categorical modeling could be an issue.

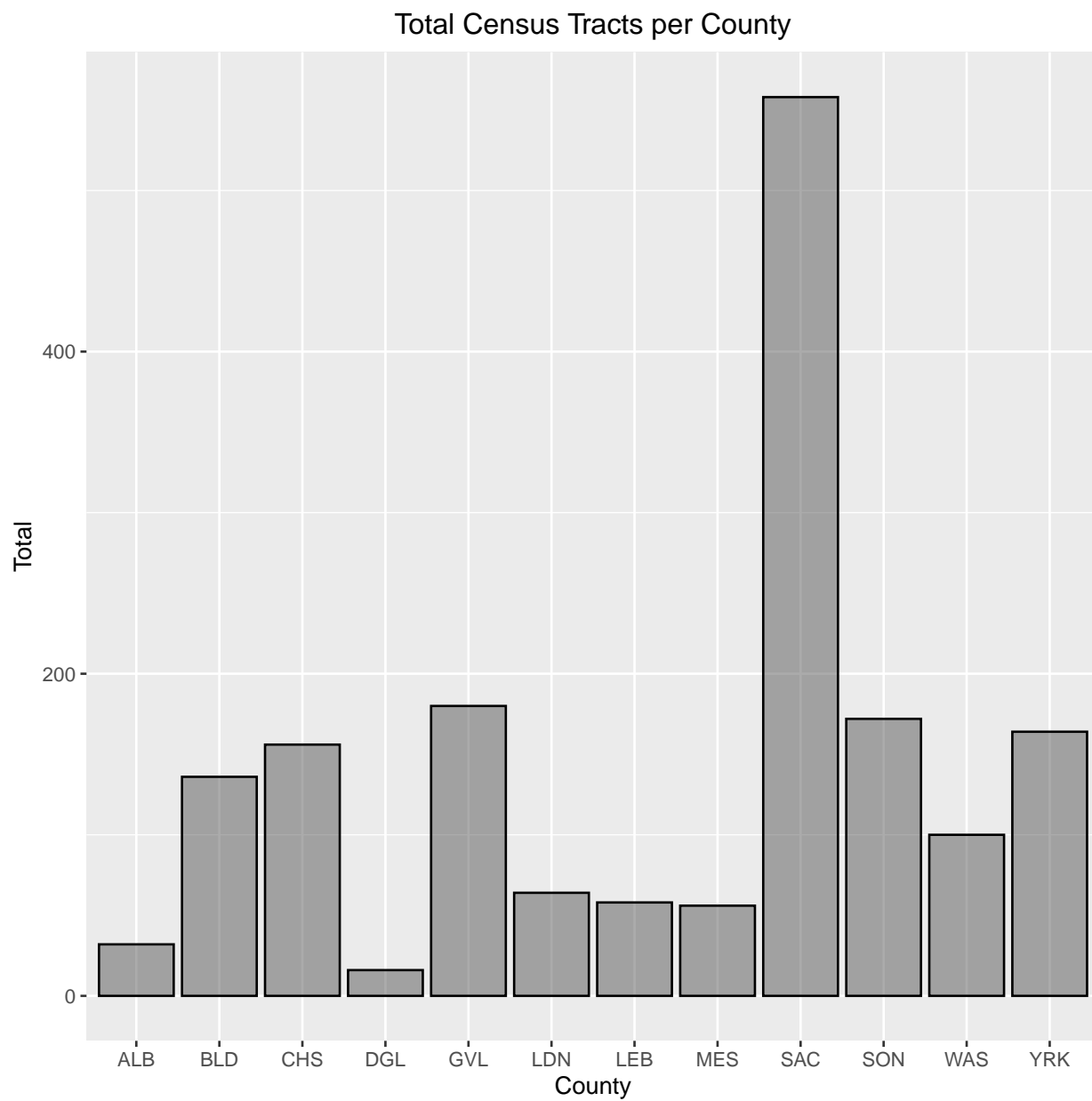


Figure 1: Total Census Tracts by County

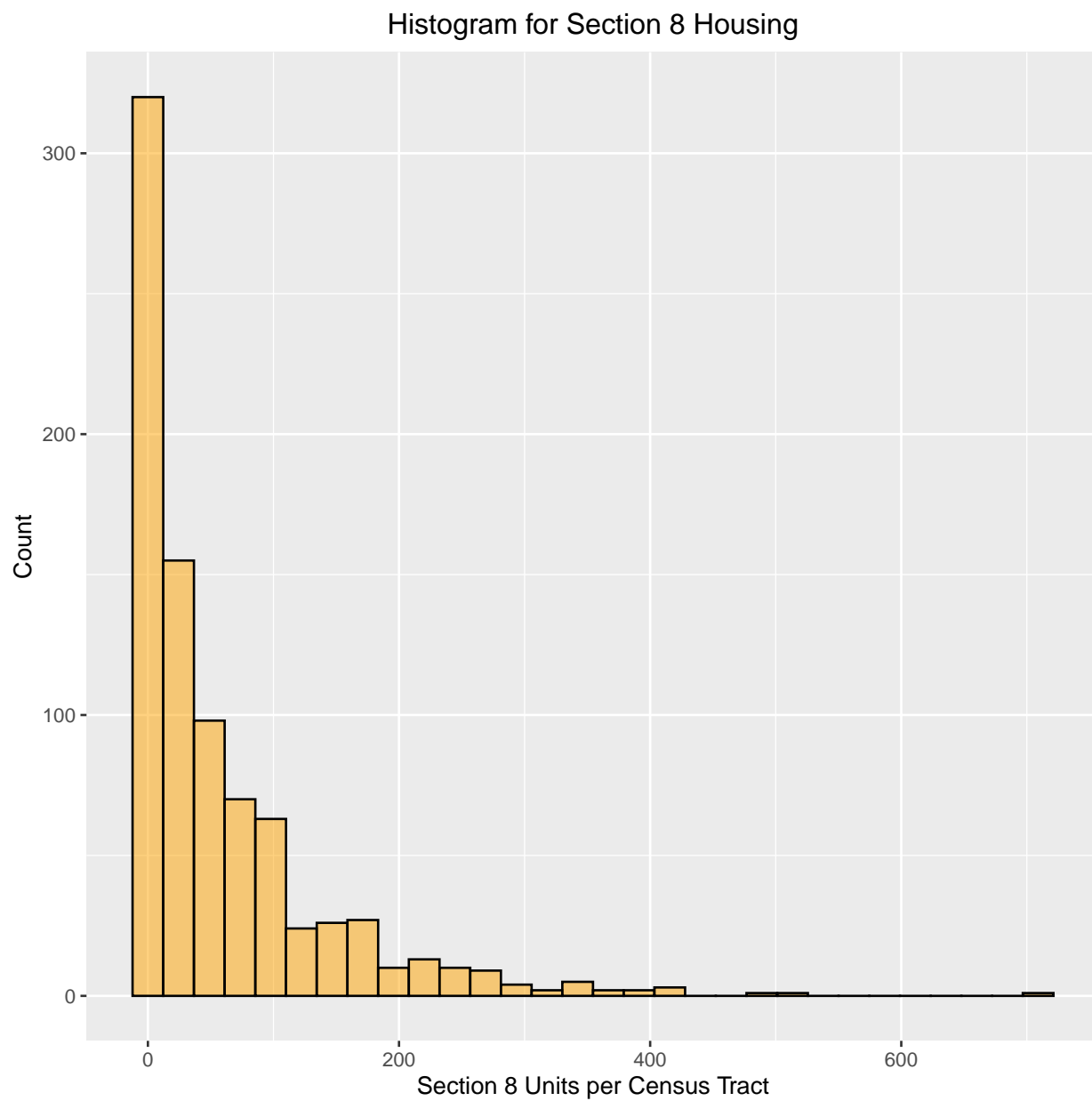


Figure 2: Section 8 Units per Census Tract

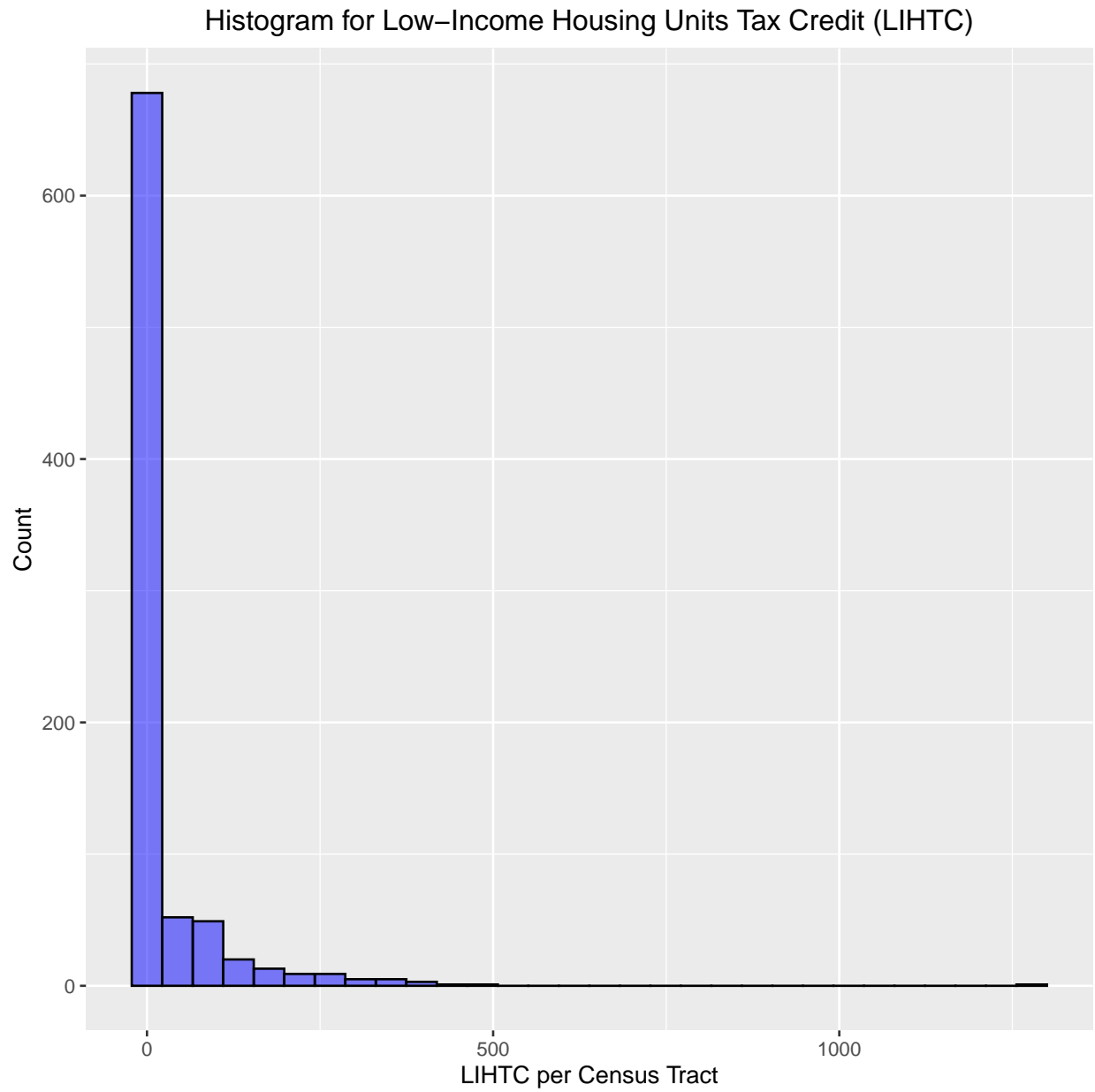


Figure 3: LIHTC Units per Census Tract

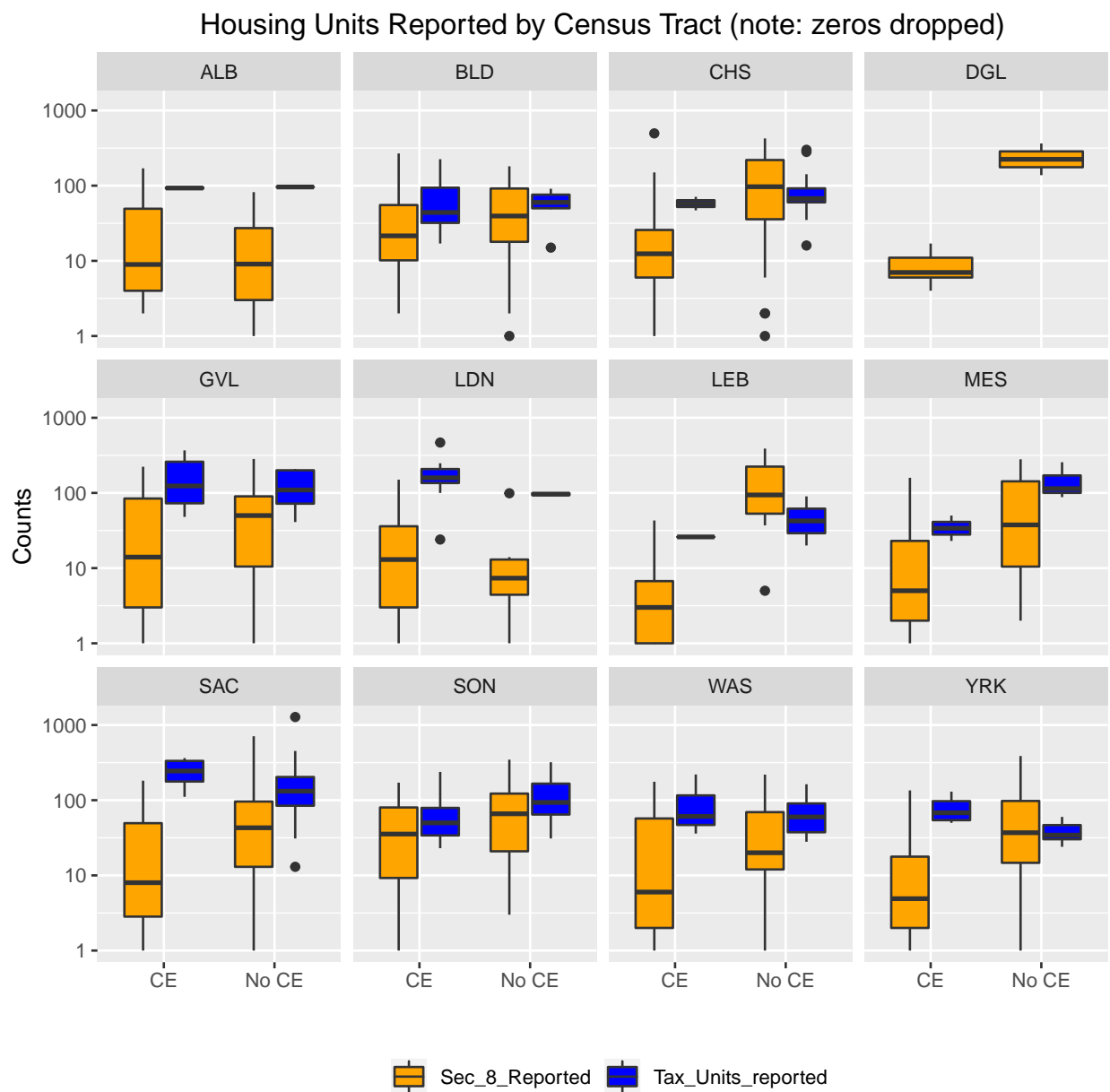


Figure 4: Housing Units Reported by Census Tract

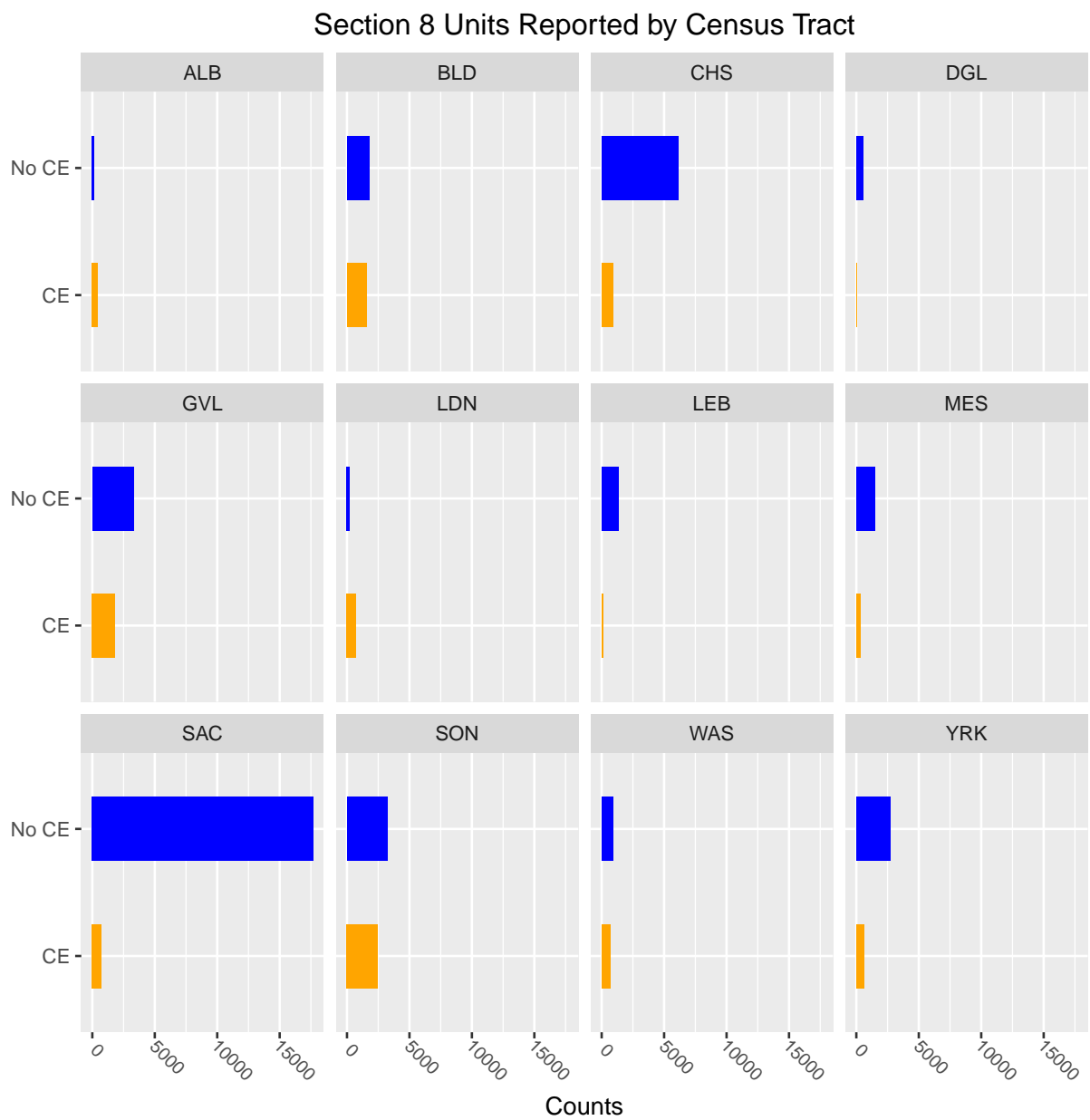


Figure 5: Section 8 Units Reported by Census Tract

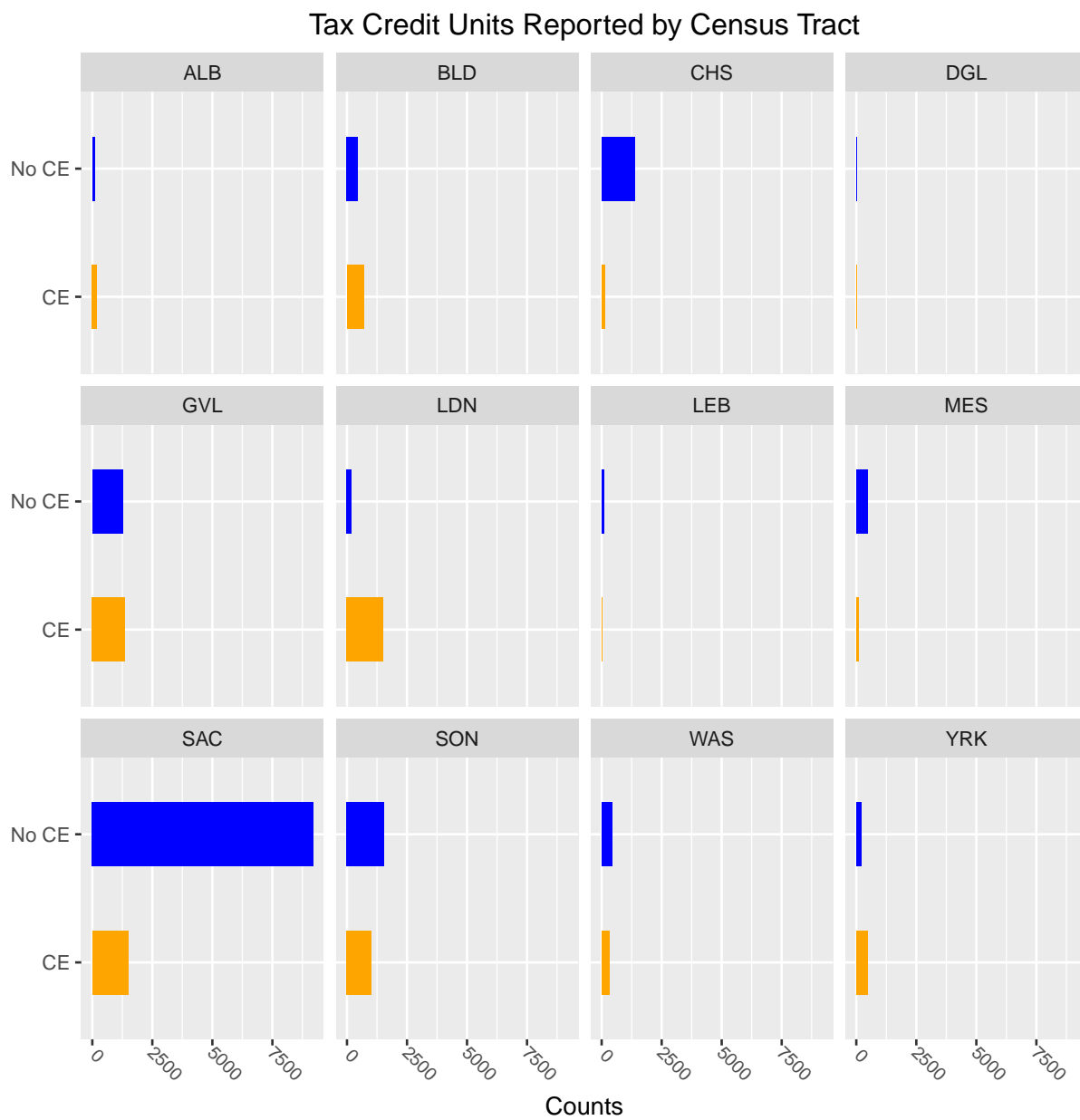


Figure 6: Tax Credit Units Reported by Census Tract

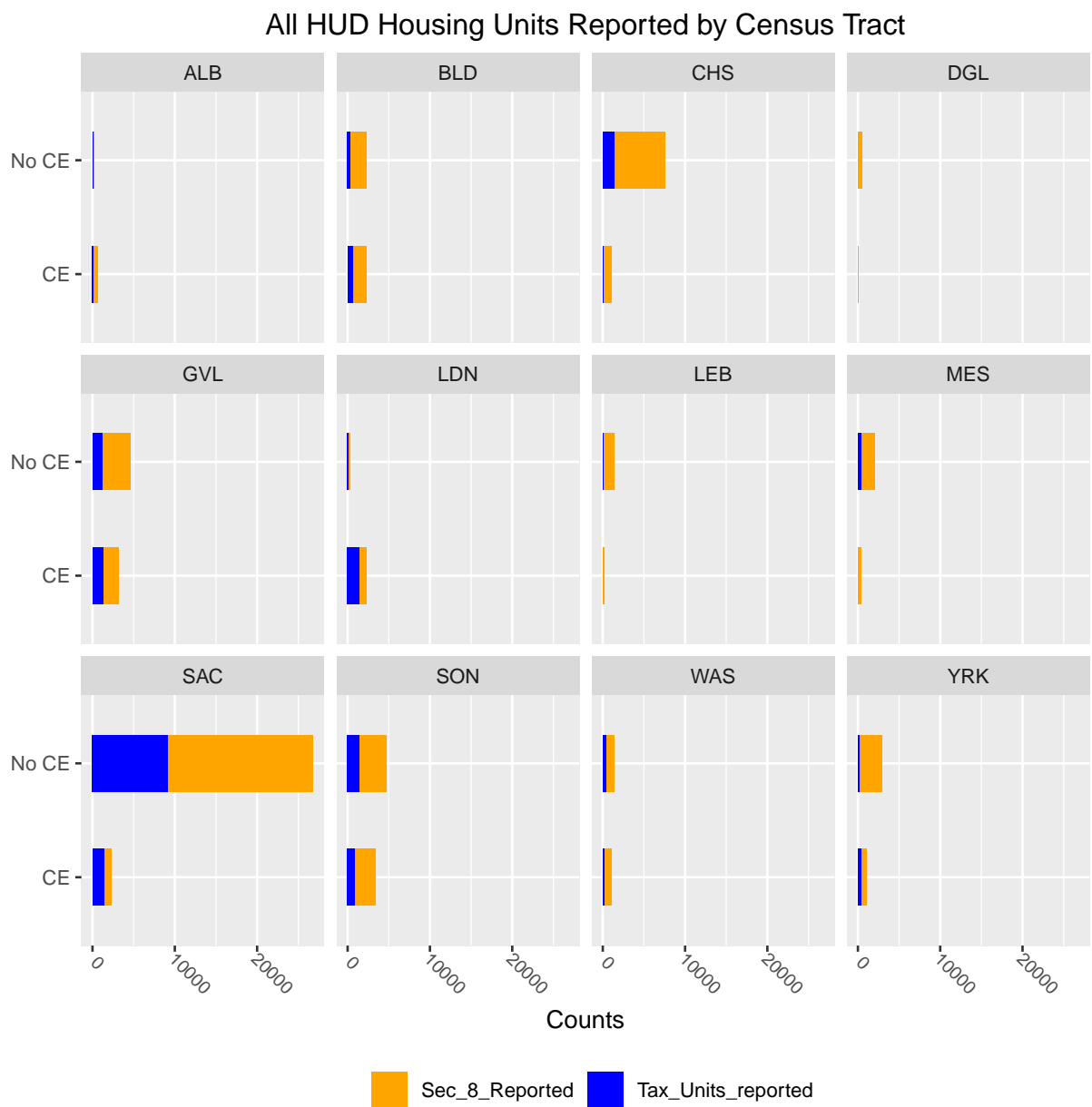


Figure 7: Housing Units Reported by Census Tract

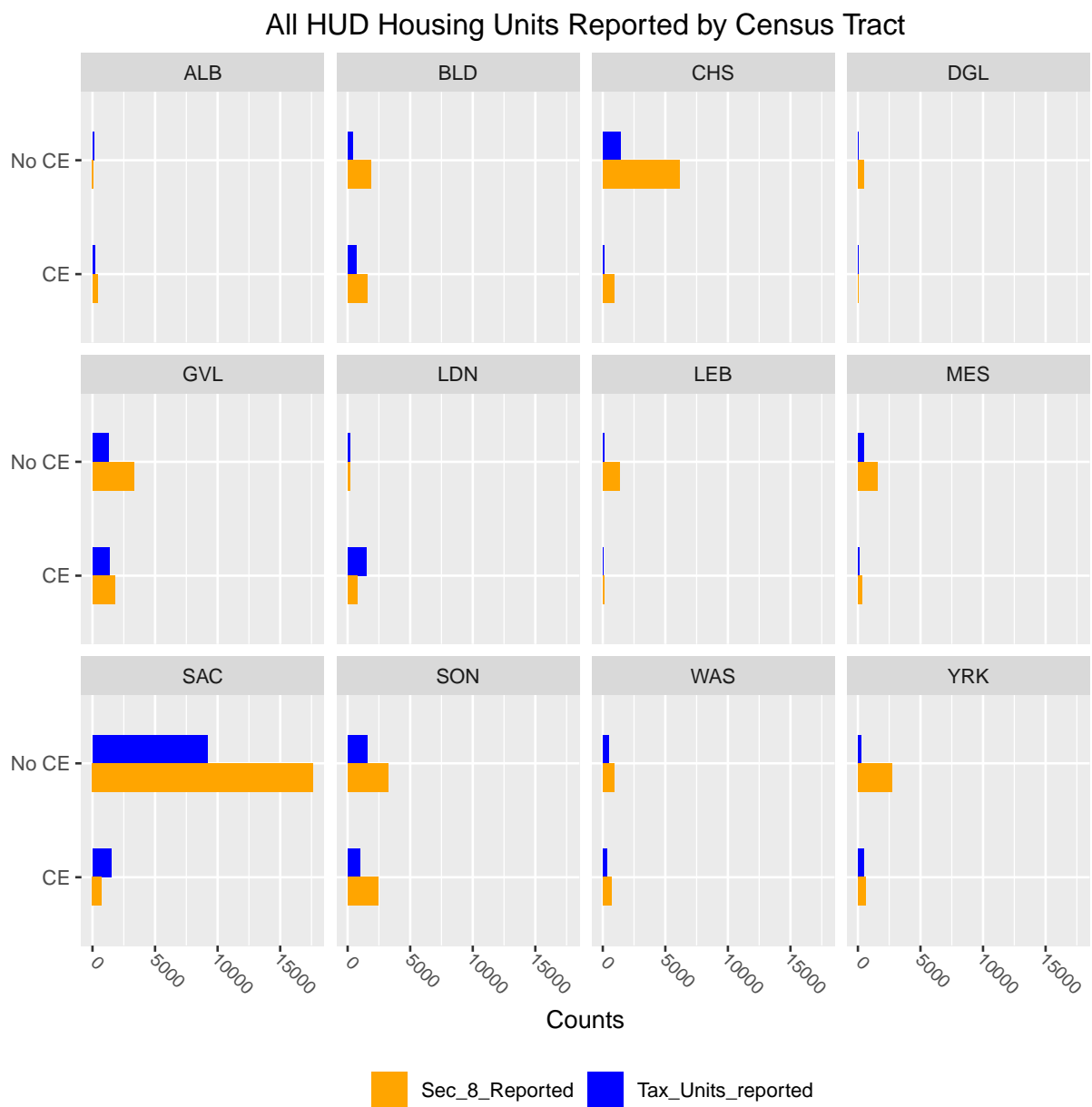


Figure 8: Housing Units Reported by Census Tract

Table 1: All Counties: Logistic Regression Results (GLM, Poisson) : HUD Housing

<i>Dependent variable:</i>	
CE_Present	
Sec_8_Reported	-0.007*** (0.001)
Tax_Units_reported	0.002** (0.001)
Constant	-0.738*** (0.070)
Observations	846
Log Likelihood	-594.702
Akaike Inf. Crit.	1,195.404
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 2: All Counties: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	22.480	25.683	4.319	13.400	31.561	40.641
Resid. Df	3	844.000	1.000	843	843.5	844.5	845
Resid. Dev	3	593.830	24.805	577.404	579.563	602.043	622.364
Pr(>Chi)	2	0.019	0.027	0.000	0.009	0.028	0.038

Table 3: All Counties: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-594.702	-617.182	44.960	0.036	0.052	0.067

Table 4: Overdispersion Test

	statistic	p.value	estimate	null.value	alternative	method	data.name
z	-11.262	1	-0.362	0	greater	Overdispersion test	model1_Poisson

Table 5: All Counties: Logistic Regression Results, Zero Inflation Model (Poisson Distribution): HUD Housing

	<i>Dependent variable:</i>
	CE_Present
Sec_8_Reported	-0.008*** (0.001)
Tax_Units_reported	0.003*** (0.001)
Constant	-0.745*** (0.070)
Observations	846
Log Likelihood	-592.625
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Table 6: BLD Regression Results: HUD Housing

<i>Dependent variable:</i>	
CE_Present	
Sec_8_Reported	-0.005 (0.004)
Tax_Units_reported	0.005 (0.005)
Constant	-0.478** (0.215)
Observations	68
Log Likelihood	-58.659
Akaike Inf. Crit.	123.318
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 7: BLD: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	0.859	0.250	0.682	0.770	0.947	1.036
Resid. Df	3	66.000	1.000	65	65.5	66.5	67
Resid. Dev	3	44.236	0.865	43.318	43.836	44.695	45.036
Pr(>Chi)	2	0.359	0.071	0.309	0.334	0.384	0.409

Table 8: BLD: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-58.659	-59.518	1.718	0.014	0.025	0.030

Table 9: BLD Overdispersion Test

	statistic	p.value	estimate	null.value	alternative	method	data.name
z	-4.550	1.000	-0.544	0	greater	Overdispersion test	BLD_model1

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

Table 10: CHS Regression Results: HUD Housing

<i>Dependent variable:</i>	
CE_Present	
Sec_8_Reported	-0.005 (0.004)
Tax_Units_reported	-0.005 (0.011)
Constant	-0.931*** (0.260)
Observations	78
Log Likelihood	-45.890
Akaike Inf. Crit.	97.780
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 11: CHS: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	2.666	3.512	0.183	1.424	3.907	5.149
Resid. Df	3	76.000	1.000	75	75.5	76.5	77
Resid. Dev	3	51.618	3.027	49.780	49.872	52.537	55.112
Pr(>Chi)	2	0.346	0.457	0.023	0.185	0.508	0.669

Table 12: CHS: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-45.890	-48.556	5.332	0.055	0.066	0.093

Table 13: BLD Overdispersion Test

	statistic	p.value	estimate	null.value	alternative	method	data.name
z	-2.778	0.997	-0.269	0	greater	Overdispersion test	CHS_model1

Table 14: GVL Regression Results: HUD Housing

	<i>Dependent variable:</i>
	CE_Present
Sec_8_Reported	-0.006 (0.004)
Tax_Units_reported	0.005 (0.004)
Constant	-0.143 (0.274)
Observations	90
Log Likelihood	-59.773
Akaike Inf. Crit.	125.546
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 15: GVL: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	1.517	0.530	1.143	1.330	1.704	1.892
Resid. Df	3	88.000	1.000	87	87.5	88.5	89
Resid. Dev	3	121.188	1.532	119.546	120.491	122.009	122.580
Pr(>Chi)	2	0.227	0.082	0.169	0.198	0.256	0.285

Table 16: GVL: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-59.773	-61.290	3.034	0.025	0.033	0.045

Table 17: LDN Regression Results: HUD Housing

	<i>Dependent variable:</i>
	CE_Present
Sec_8_Reported	0.003 (0.014)
Tax_Units_reported	0.007 (0.008)
Constant	0.463 (0.455)
Observations	32
Log Likelihood	-18.791
Akaike Inf. Crit.	43.583
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 18: LDN: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	1.083	0.126	0.994	1.039	1.128	1.172
Resid. Df	3	30.000	1.000	29	29.5	30.5	31
Resid. Dev	3	38.636	1.085	37.583	38.080	39.163	39.750
Pr(>Chi)	2	0.299	0.028	0.279	0.289	0.309	0.319

Table 19: LDN: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-18.791	-19.875	2.167	0.055	0.065	0.092

Table 20: LEB Regression Results: HUD Housing

	<i>Dependent variable:</i>
	CE_Present
Sec_8_Reported	-0.090** (0.038)
Tax_Units_reported	0.137 (0.285)
Constant	3.130*** (1.089)
Observations	29
Log Likelihood	-5.563
Akaike Inf. Crit.	17.126
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 21: LEB: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	12.399	17.160	0.265	6.332	18.466	24.533
Resid. Df	3	27.000	1.000	26	26.5	27.5	28
Resid. Dev	3	19.480	14.241	11.126	11.258	23.657	35.924
Pr(>Chi)	2	0.303	0.429	0.00000	0.152	0.455	0.607

Table 22: LEB: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-5.563	-17.962	24.798	0.690	0.575	0.809

Table 23: MES Regression Results: HUD Housing

	<i>Dependent variable:</i>
	CE_Present
Sec_8_Reported	-0.013 (0.010)
Tax_Units_reported	0.004 (0.016)
Constant	0.018 (0.513)
Observations	28
Log Likelihood	-16.506
Akaike Inf. Crit.	39.012
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 24: MES: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	1.743	2.357	0.077	0.910	2.576	3.409
Resid. Df	3	26.000	1.000	25	25.5	26.5	27
Resid. Dev	3	34.200	1.991	33.012	33.050	34.794	36.498
Pr(>Chi)	2	0.423	0.507	0.065	0.244	0.602	0.782

Table 25: MES: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-16.506	-18.249	3.486	0.096	0.117	0.161

Table 26: SAC Regression Results: HUD Housing

	<i>Dependent variable:</i>
	CE_Present
Sec_8_Reported	-0.017** (0.007)
Tax_Units_reported	0.004** (0.002)
Constant	-1.834*** (0.289)
Observations	279
Log Likelihood	-73.232
Akaike Inf. Crit.	152.464
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 27: SAC: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	6.195	0.609	5.764	5.979	6.410	6.625
Resid. Df	3	277.000	1.000	276	276.5	277.5	278
Resid. Dev	3	152.515	6.199	146.464	149.346	155.540	158.853
Pr(>Chi)	2	0.013	0.004	0.010	0.012	0.015	0.016

Table 28: SAC: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-73.232	-79.426	12.389	0.078	0.043	0.100

Table 29: SON Regression Results: HUD Housing

<i>Dependent variable:</i>	
CE_Present	
Sec_8_Reported	-0.007* (0.004)
Tax_Units_reported	-0.001 (0.004)
Constant	0.694** (0.324)
Observations	86
Log Likelihood	-56.690
Akaike Inf. Crit.	119.379
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 30: SON: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	2.548	3.524	0.056	1.302	3.794	5.040
Resid. Df	3	84.000	1.000	83	83.5	84.5	85
Resid. Dev	3	115.097	2.926	113.379	113.408	115.956	118.476
Pr(>Chi)	2	0.419	0.557	0.025	0.222	0.616	0.812

Table 31: SON: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-56.690	-59.238	5.097	0.043	0.058	0.077

Table 32: WAS Regression Results: HUD Housing

	<i>Dependent variable:</i>
	CE_Present
Sec_8_Reported	-0.006 (0.007)
Tax_Units_reported	-0.0001 (0.009)
Constant	0.358 (0.342)
Observations	50
Log Likelihood	-33.927
Akaike Inf. Crit.	73.854
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 33: WAS: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	0.570	0.806	0.0001	0.285	0.855	1.140
Resid. Df	3	48.000	1.000	47	47.5	48.5	49
Resid. Dev	3	68.234	0.658	67.854	67.854	68.424	68.994
Pr(>Chi)	2	0.639	0.500	0.286	0.462	0.816	0.992

Table 34: WAS: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-33.927	-34.497	1.140	0.017	0.023	0.030

Table 35: YRK Regression Results: HUD Housing

	<i>Dependent variable:</i>
	CE_Present
Sec_8_Reported	-0.033*** (0.011)
Tax_Units_reported	0.035** (0.016)
Constant	0.726** (0.304)
Observations	82
Log Likelihood	-45.494
Akaike Inf. Crit.	96.988
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 36: YRK: Analysis of Deviance

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	1.000	0.000	1.000	1.000	1.000	1.000
Deviance	2	11.246	6.444	6.690	8.968	13.525	15.803
Resid. Df	3	80.000	1.000	79	79.5	80.5	81
Resid. Dev	3	100.716	11.550	90.988	94.333	105.580	113.481
Pr(>Chi)	2	0.005	0.007	0.0001	0.002	0.007	0.010

Table 37: YRK: McFadden Statistic:similar to R2

llh	llhNull	G2	McFadden	r2ML	r2CU
-45.494	-56.740	22.493	0.198	0.240	0.320