An Evaluation of Population Estimates

Nazrul Hoque
Texas State Data Center
University of Texas at San Antonio

Ratio-Correlation Method

- **The Ratio-Correlation Method is a multiple regression-based technique which compares change in subarea to change occurring in super area. Such estimates are developed using the following multiple regression equation:**

$$y = B_0 + B_1 X_1 + B_2 X_2 + \ldots\ldots B_n X_n + e$$

where, y = **the dependent variable to be estimated (population)**

B_0 = **the intercept to be estimated**

B_i = **the coefficient to be estimated**

X_i = **indepdent variables, such as births, deaths, etc.**

e = error term

Ratio-Correlation Method (cont)

Step 1. **Obtain Coefficients for Change Occurring at Previous Points in Time (such as 1990-2000)**

$$\frac{\text{(County Pop, 2000)/State Pop, 2000)}}{\text{(County pop, 1990)/State Pop, 1990)}} = \alpha_0$$

$$+B_1 \quad \frac{\text{(County Births, 2000)/(State Births, 2000)}}{\text{(County Births, 1990)/(State Births, 1990)}}$$

$$+B_2 \quad \frac{\text{(County Deaths, 2000)/(State Deaths 2000)}}{\text{(County Deaths, 1990)/(State Deaths, 1990)}}$$

$$+B_3 \quad \frac{\text{(County Sch. Enr., 2000)/(State Sch. Enr., 2000)}}{\text{(County Sch. Enr., 1990)/(State Sch. Enr., 1990)}}$$

Step 2. **Apply Coefficients to Change During Recent Time Period**

Housing Unit Method

$P_t = (OHU_t \times PPH_t) + GQ_t$

Where;

P_t = Total Population at the time of estimate

OHU_t = Occupied housing units at the time of estimate

GQ_t = Group Quarters population at the time of estimate

PPH_t = Household size or population perhousehold at the time of estimate (P_t / H_t)

P_t = Non Group quarters population at the time of estimate

H_t = Non Group quarters Occupied Households at the time of estimate

Housing Unit Method

$P_t = (OHU_o + BP_t - DP_t - VU_t) \times PPH_t + GQ_t$

Where;

P_t = Total Population at the time of estimate

OHU_o = Occupied housing units on the last census date

GQ_t = Group Quarters population at the time of estimate

PPH_o = Household size or population per household at the last census date (P_o / H_o)

P_o = Non Group quarters population at the last census

H_o = Non Group quarters Occupied Households at the last census

BP_t = Building permits issued between P_t and P_o

DP_t = Demolition permits issued between P_t and P_o

VU_t = Vacant Unit between P_t and P_o

Estimating Populations Using Component Methods

Base population

- **+ Births**
- **- Deaths**
- **+ Net Migration**

Estimate Error

Estimate error may be defined as the difference between the estimated population for a particular area in a particular year and the actual population for the same area and year:

$$\text{Estimation Error} = \text{Estimate Value} - \text{Census Count}$$

$$\text{Percent Error} = \frac{\text{Estimate Value - Census Count}}{\text{Census Count}} \times 100$$

Error rates were examined by:

- **Population Size**
- **Percent Population Change**
- **Range of Error**

Commonly Used Measures of Error

$$\text{Mean Percent Error (MPE)} = \frac{\sum_{i=1}^{n}\left(\frac{\text{Estimated Value - Census Count}}{\text{Census Count}}\right)}{n} \times 100$$

$$\text{Mean Absolute Percent Error (MAPE)} = \frac{\sum_{i=1}^{n}\left|\frac{\text{Estimated Value - Census Count}}{\text{Census Count}}\right|}{n} \times 100$$

$$\text{Mean Percent Absolute Difference (MPAD)} = \frac{\sum_{i=1}^{n}\left|\text{Estimated Value - Census Count}\right|}{\sum_{i=1}^{n}\text{Census Count}} \times 100$$

Example of Use of Error Measures

County	Estimated Population 2000	Census Counts 2000	Error (difference)	Percent Error	Absolute Error	Absolute Percent Error
County – 1	15,113	15,304	-191	-1.3	191	1.3
County – 2	18,204	17,496	708	4.1	708	4.1
County – 3	16,258	15,683	575	3.7	575	3.7
County – 4	27,774	27,392	382	1.4	382	1.4
County – 5	21,561	20,796	765	3.7	765	3.7
County – 6	195,193	191,177	3,446	1.8	3,446	1.8
Total	294,103	288,418	5,685	13.4	6,067	16.0

$$\textbf{Mean Percent Error (MPE)} = \frac{\sum_{i=1}^{n}\left(\frac{\textbf{Estimate Value - Census Count}}{\textbf{Census Count}}\right)\times 100}{n} = \frac{13.4}{6} = 2.2$$

Example of Use of Error Measures (cont)

County	Estimated Population 2000	Census Counts 2000	Error (difference)	Percent Error	Absolute Error	Absolute Percent Error
County – 1	15,113	16,304	-1,191	-7.30	1,191	7.30
County – 2	17,324	17,496	-172	-0.98	172	0.98
County – 3	15,559	15,683	-124	-0.79	124	0.79
County – 4	27,734	27,392	342	1.24	342	1.24
County – 5	22,010	20,805	1,205	5.79	1,205	5.79
County – 6	195,660	191,747	3,913	2.04	3,913	2.04
Total	293,400	289,427	3,973	0.00	6,947	18.14

$$\textbf{Mean Percent Error (MPE)} = \frac{\sum_{i=1}^{n}\left(\frac{\textbf{Estimate Value - Census Count}}{\textbf{Census Count}}\right)\times 100}{n} = \frac{0.0}{6} = 0.0$$

Example of Use of Error Measures (cont)

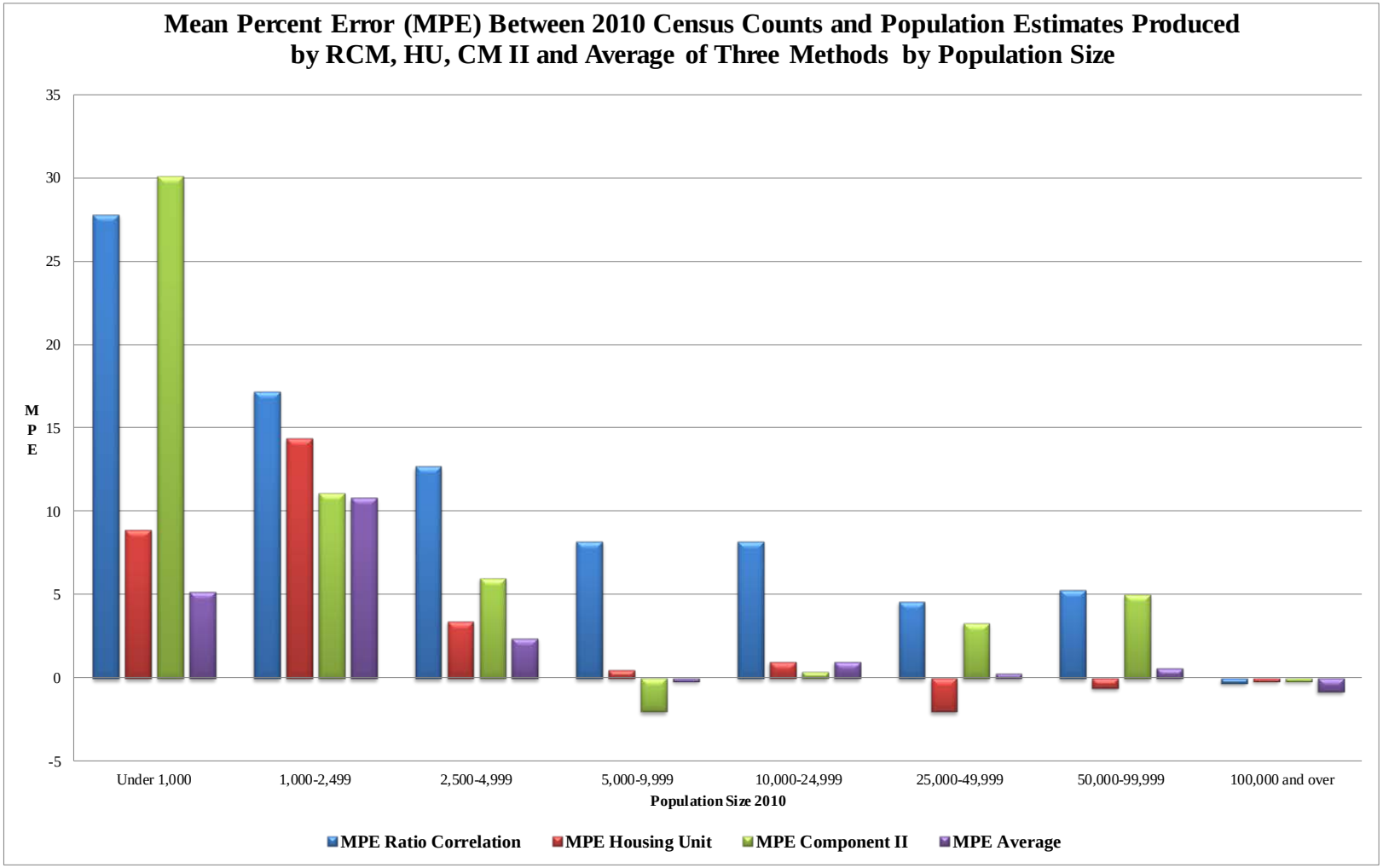
County	Estimated Population 2000	Census Counts 2000	Error (difference)	Percent Error	Absolute Error	Absolute Percent Error
County – 1	15,113	15,304	-191	-1.3	191	1.3
County – 2	18,204	17,496	708	4.1	708	4.1
County – 3	16,258	15,683	575	3.7	575	3.7
County – 4	27,774	27,392	382	1.4	382	1.4
County – 5	21,561	20,796	765	3.7	765	3.7
County – 6	195,193	191,177	3,446	1.8	3,446	1.8
Total	294,103	288,418	5,685	13.4	6,067	16.0

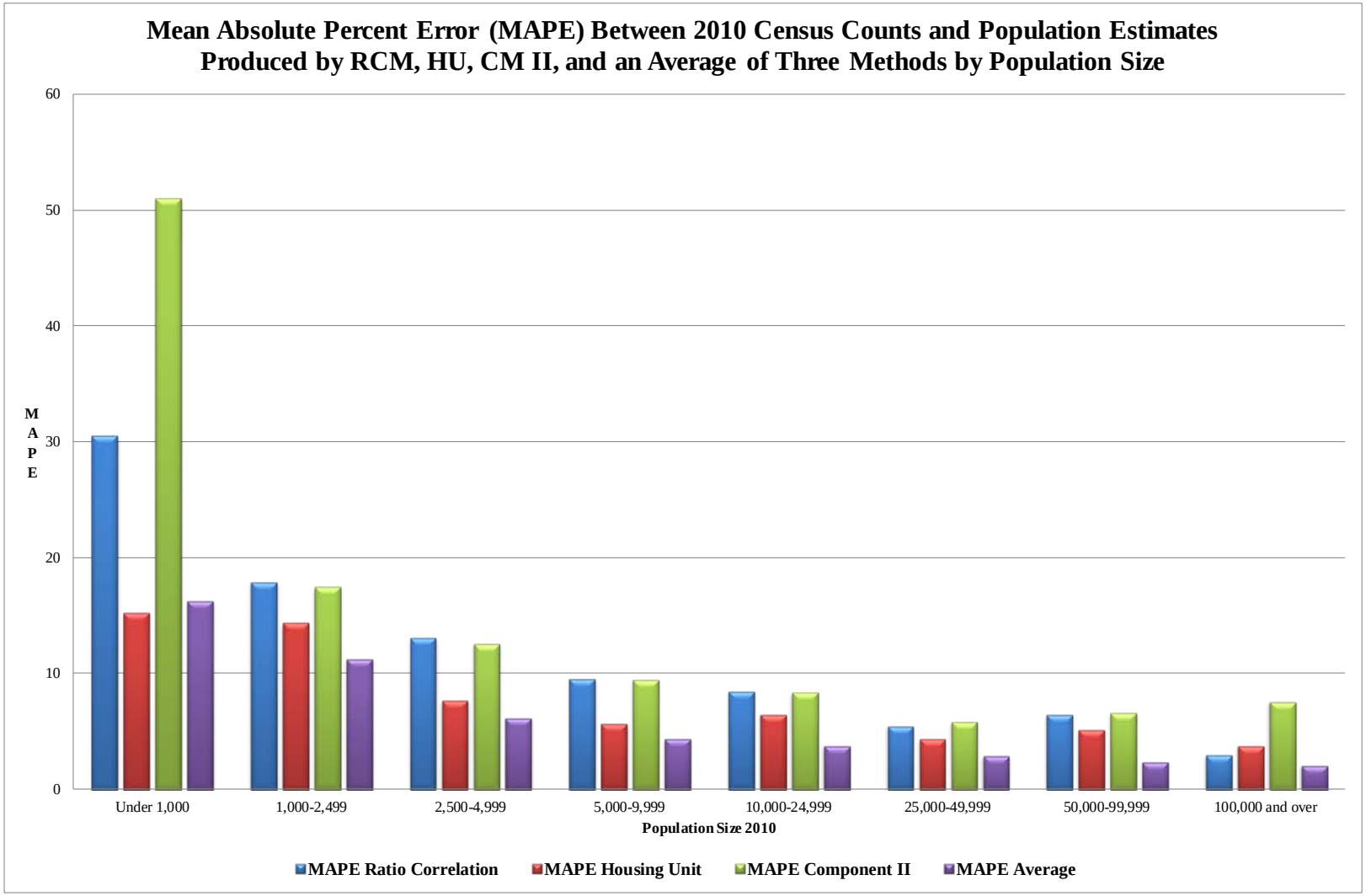
$$\textbf{Mean Absolute Percent Error (MAPE)} = \frac{\sum_{i=1}^{n}\left|\frac{\textbf{Estimate Value - Census Count}}{\textbf{Census Count}}\right| \times 100}{n} = \frac{16.0}{6} = 2.7$$

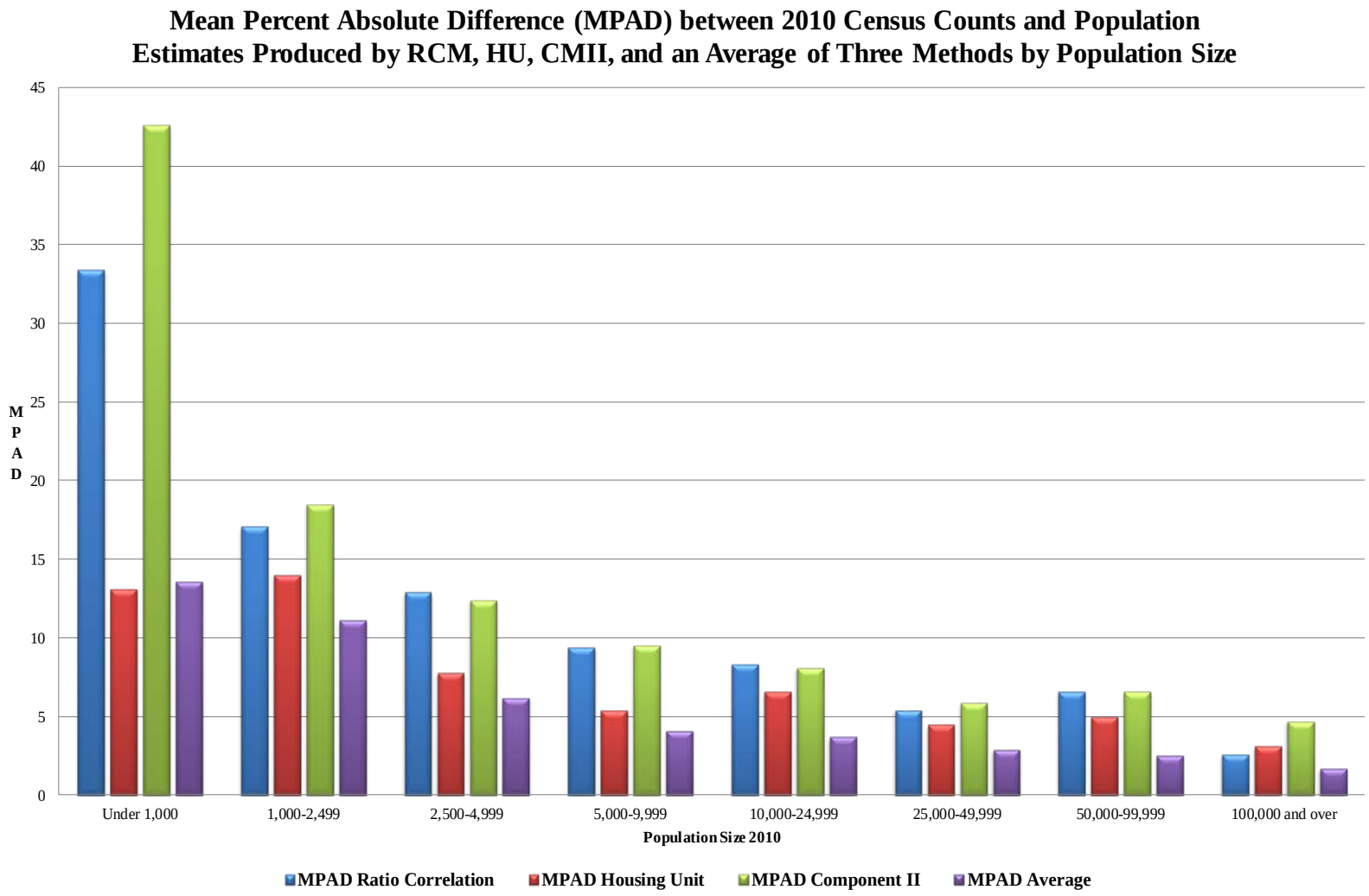
Example of Use of Error Measures (cont)

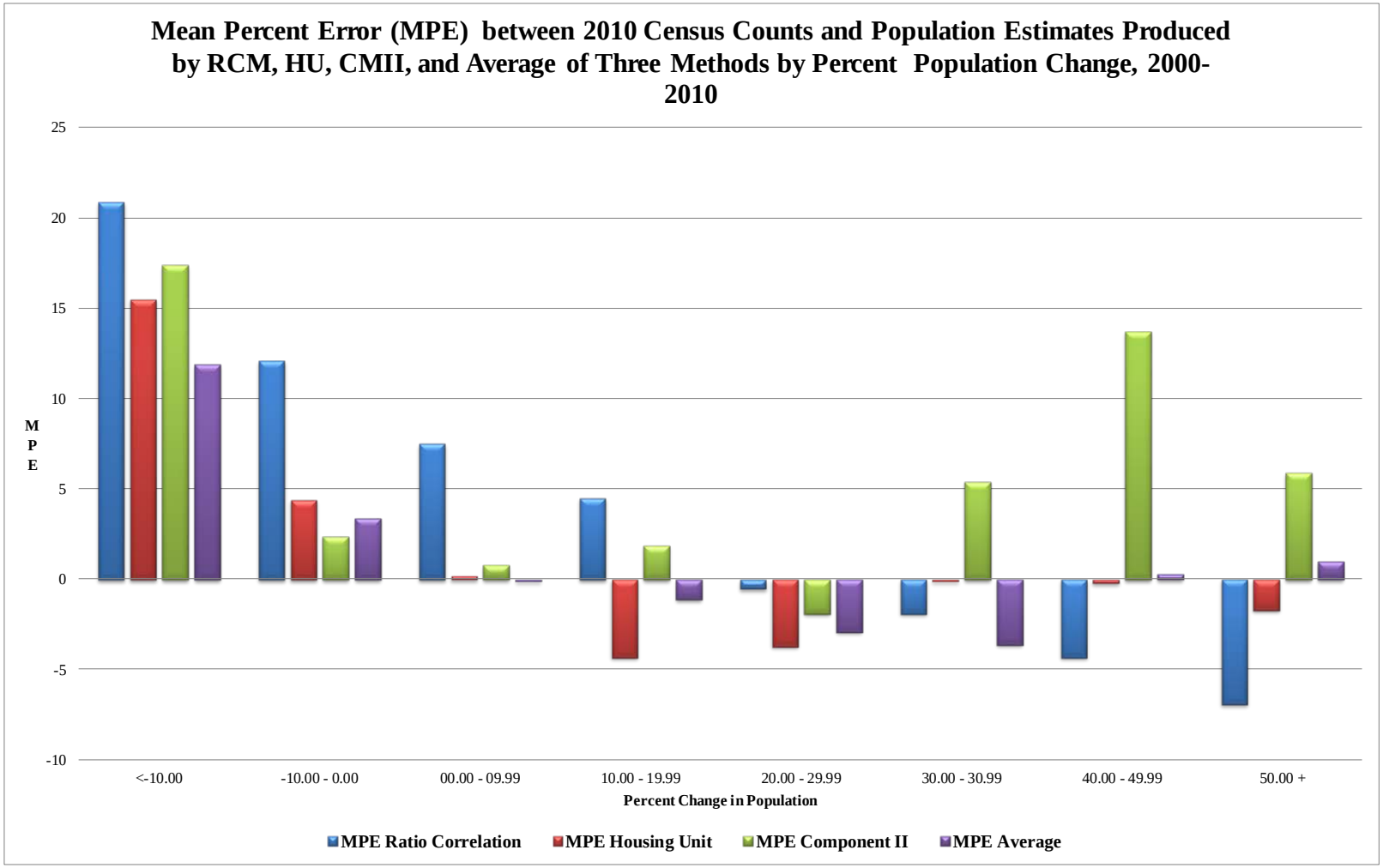
County	Estimated Population 2000	Census Counts 2000	Error (difference)	Percent Error	Absolute Error	Absolute Percent Error
County – 1	15,113	15,304	-191	-1.3	191	1.3
County – 2	18,204	17,496	708	4.1	708	4.1
County – 3	16,258	15,683	575	3.7	575	3.7
County – 4	27,774	27,392	382	1.4	382	1.4
County – 5	21,561	20,796	765	3.7	765	3.7
County – 6	195,193	191,177	3,446	1.8	3,446	1.8
Total	294,103	288,418	5,685	13.4	6,067	16.0

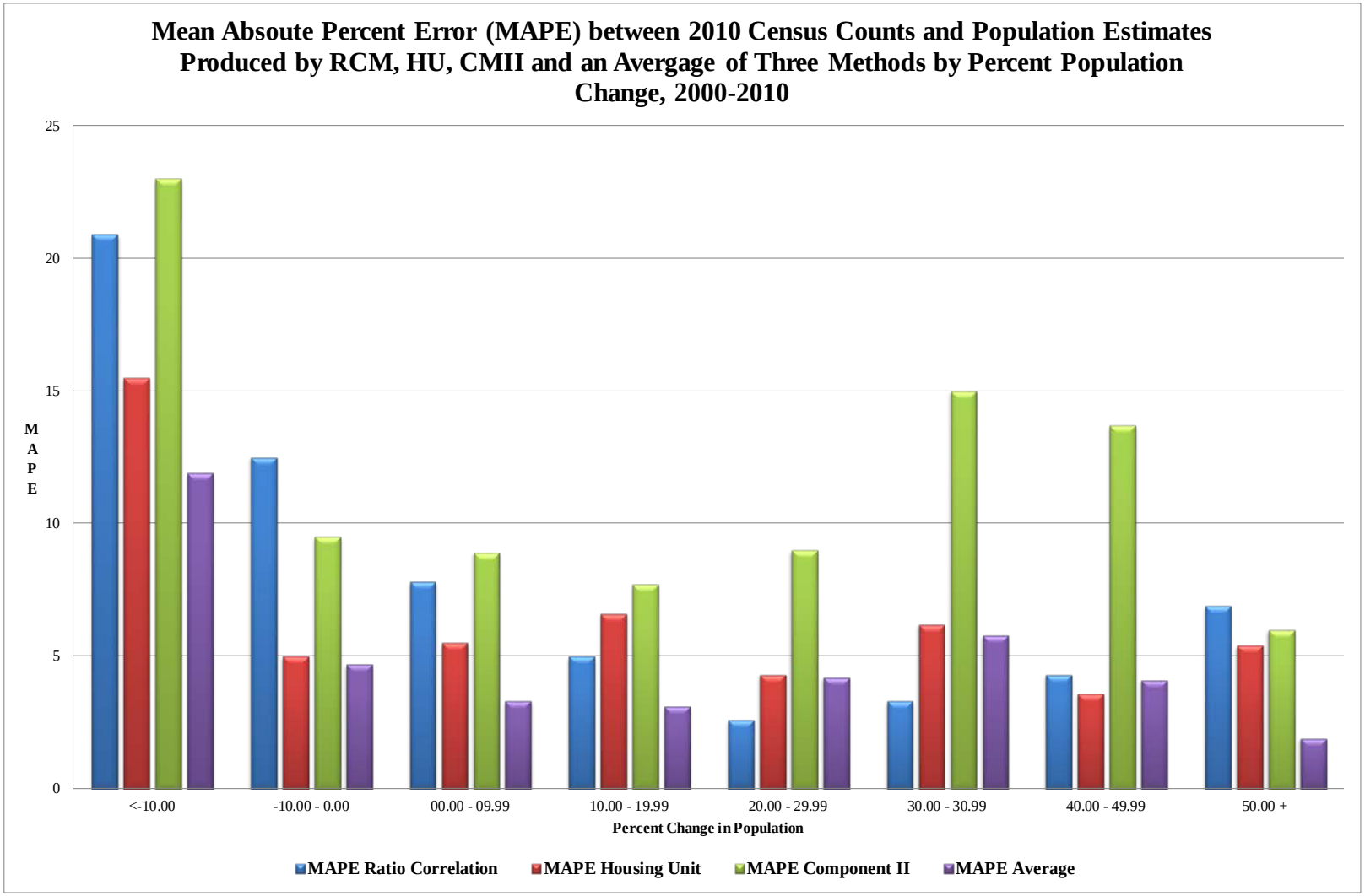
$$\textbf{Mean Percent Absolute Difference (MPAD)} = \frac{\sum_{i=1}^{n} |\textbf{Estimate Value - Census Count}|}{\sum_{i=1}^{n} \textbf{Census Count}} \times 100 = \frac{6,067}{288,418} = 2.1$$

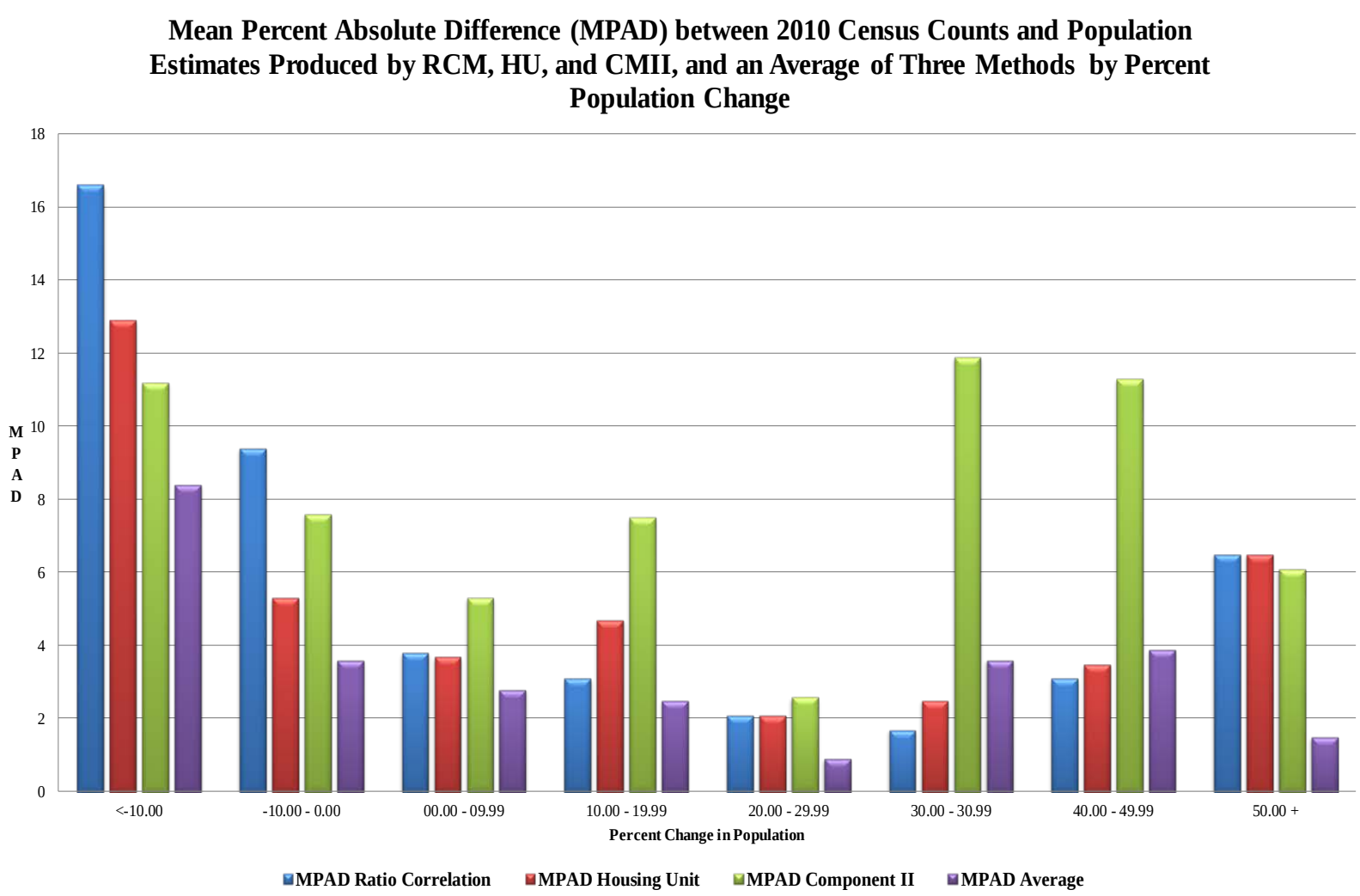

Mean Percent Error (MPE) Between 2010 Census Counts and Population Estimates Produced by RCM, HU, CM II and Average of Three Methods by Population Size
35
30
25
20
15
10
5
0
-5
MPE
Under 1,000
1,000-2,499
2,500-4,999
5,000-9,999
10,000-24,999
25,000-49,999
50,000-99,999
100,000 and over
Population Size 2010
MPE Ratio Correlation
MPE Housing Unit
MPE Component II
MPE Average

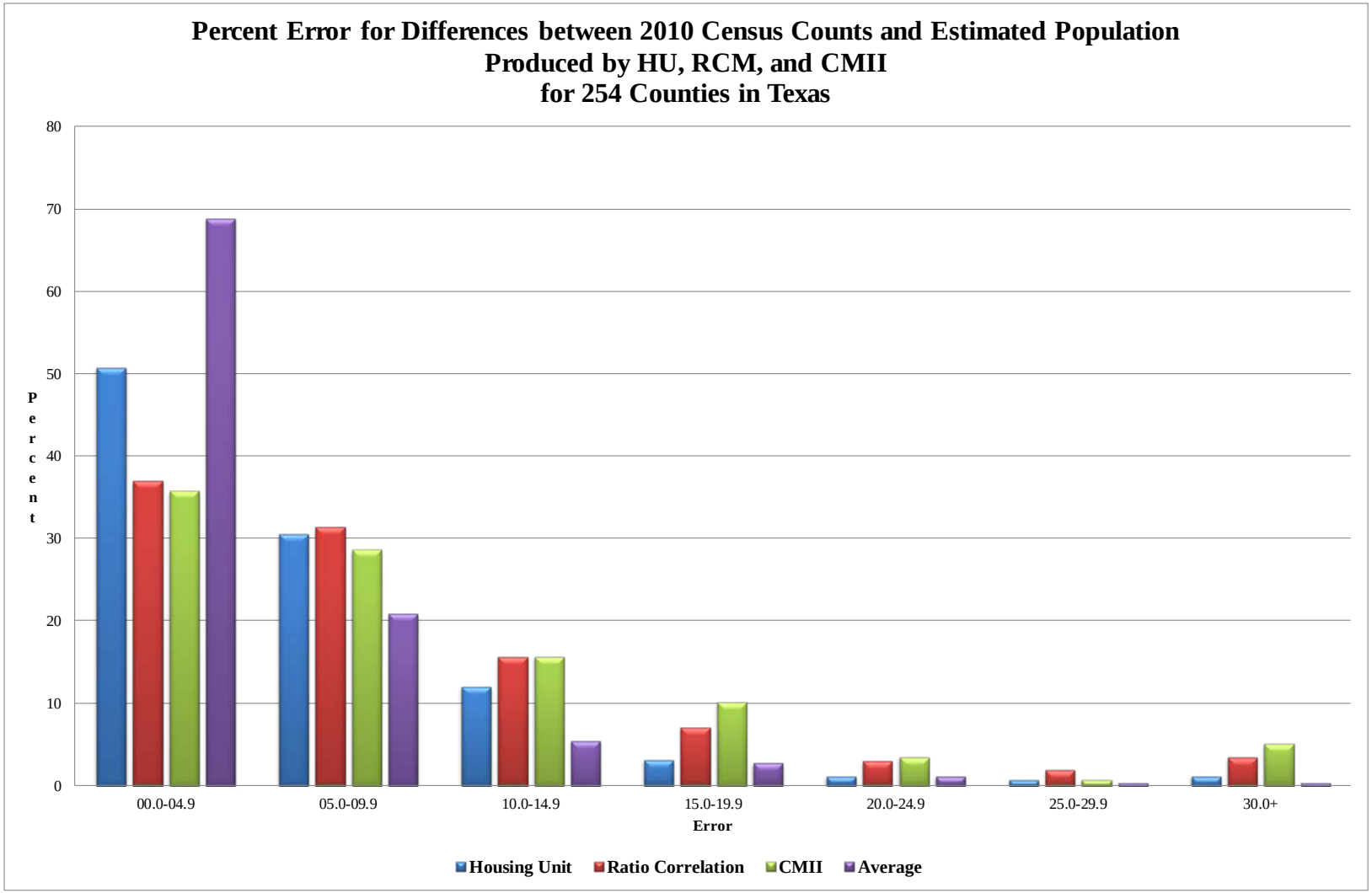

Mean Absolute Percent Error (MAPE) Between 2010 Census Counts and Population Estimates Produced by RCM, HU, CM II, and an Average of Three Methods by Population Size
MAPE
60
50
40
30
20
10
0
Under 1,000
1,000-2,499
2,500-4,999
5,000-9,999
10,000-24,999
25,000-49,999
50,000-99,999
100,000 and over
Population Size 2010
MAPE Ratio Correlation
MAPE Housing Unit
MAPE Component II
MAPE Average


Mean Percent Absolute Difference (MPAD) between 2010 Census Counts and Population Estimates Produced by RCM, HU, CMII, and an Average of Three Methods by Population Size
MPAD
45
40
35
30
25
20
15
10
5
0
Under 1,000
1,000-2,499
2,500-4,999
5,000-9,999
10,000-24,999
25,000-49,999
50,000-99,999
100,000 and over
Population Size 2010
MPAD Ratio Correlation
MPAD Housing Unit
MPAD Component II
MPAD Average


Mean Percent Error (MPE) between 2010 Census Counts and Population Estimates Produced by RCM, HU, CMII, and Average of Three Methods by Percent Population Change, 2000-2010
25
20
15
10
5
0
-5
-10
MPE
<-10.00
-10.00 - 0.00
00.00 - 09.99
10.00 - 19.99
20.00 - 29.99
30.00 - 30.99
40.00 - 49.99
50.00 +
Percent Change in Population
MPE Ratio Correlation
MPE Housing Unit
MPE Component II
MPE Average


Mean Absoute Percent Error (MAPE) between 2010 Census Counts and Population Estimates Produced by RCM, HU, CMII and an Avergage of Three Methods by Percent Population Change, 2000-2010
25
20
15
10
5
0
MAPE
<-10.00
-10.00 - 0.00
00.00 - 09.99
10.00 - 19.99
20.00 - 29.99
30.00 - 30.99
40.00 - 49.99
50.00 +
Percent Change in Population
MAPE Ratio Correlation
MAPE Housing Unit
MAPE Component II
MAPE Average


Mean Percent Absolute Difference (MPAD) between 2010 Census Counts and Population Estimates Produced by RCM, HU, and CMII, and an Average of Three Methods by Percent Population Change
MPAD
18
16
14
12
10
8
6
4
2
0
<-10.00
-10.00 - 0.00
00.00 - 09.99
10.00 - 19.99
20.00 - 29.99
30.00 - 30.99
40.00 - 49.99
50.00 +
Percent Change in Population
MPAD Ratio Correlation
MPAD Housing Unit
MPAD Component II
MPAD Average


Percent Error for Differences between 2010 Census Counts and Estimated Population
Produced by HU, RCM, and CMII
for 254 Counties in Texas
80
70
60
50
40
30
20
10
0
Percent
00.0-04.9
05.0-09.9
10.0-14.9
15.0-19.9
20.0-24.9
25.0-29.9
30.0+
Error
Housing Unit
Ratio Correlation
CMII
Average

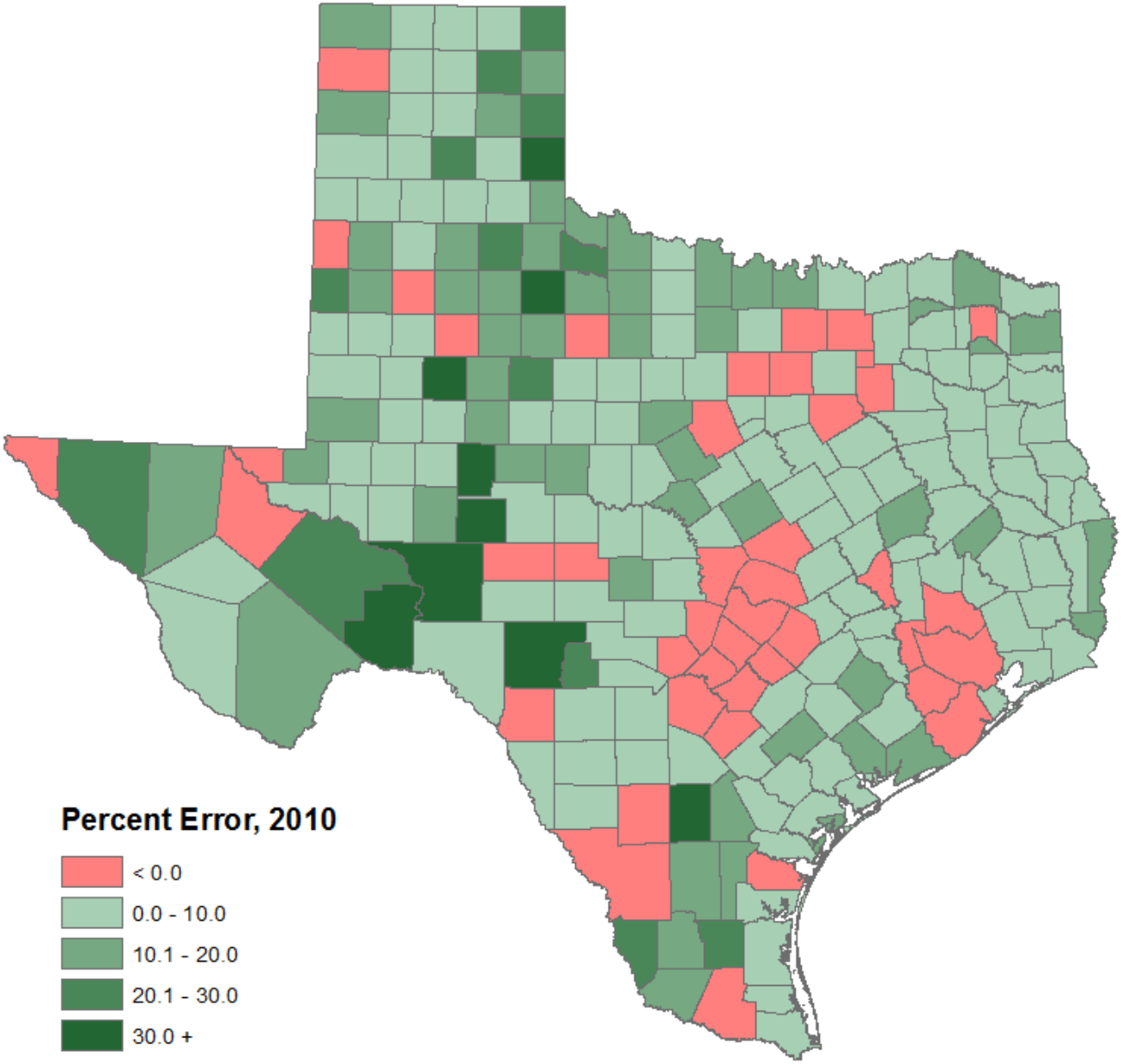



Figure 3: Percent Error from Comparison of Population Estimates Produced by Ratio-Correlation Method and 2010 Census Counts