RESEARCH REPORT SERIES
(Statistics #2013-05)

A Visual Proof, a Test, and an Extension of a Simple Tool for Comparing Competing Estimates

Tommy Wright

Center for Statistical Research & Methodology
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: September 10, 2013
Revised and reissued: February 24, 2014

A Visual Proof, a Test, and an Extension of A Simple Tool for Comparing Competing Estimates

Tommy Wright
U. S. Bureau of the Census

Abstract

A common practice of a national statistical agency (and most scientific investigations) is to assess its estimate for a particular real-valued characteristic (parameter) by comparing it to competing estimates from other sources. The nature of the assessment is to explain the differences. While assessing the quality of energy estimates (e.g., reserves, production, supply, sales, consumption,...), Tsao and Wright (1983) introduced a simple tool (maximum ratio) and used it to prove a statement regarding the closeness of a set containing a particular estimate and competing estimates to the unknown true value of the targeted real-valued parameter. In this paper, we (1) provide a visual proof of the (*Main Result*), (2) present a test of unacceptability for at least one of the competing estimates using the maximum ratio, (3) give an application assessing the national count from the 2010 U.S. Census, and (4) extend the maximum ratio and statement to vector-valued parameters. The test can be applied in any situation where several measurements of the same phenomenon exist and one wants to quantify with certainty how far away from the truth at least one of the estimates is. Even if all estimates are near each other, this does not necessarily imply that they are near the unknown true value. The test sends a signal calling for further investigation when a pre-set standard is exceeded.

Key Words: Accuracy, Maximum Ratio, Test of Unacceptability.

1. INTRODUCTION: COMPARISONS

1.1 The Utility of Comparisions

One basic step toward improving the quality of estimates produced by a national statistical agency is to identify problem areas that require special attention. Comparisons with competing estimates from other sources, when possible, can serve to expose existing problems for one or more data sources.

If θ is some parameter (such as a total or a mean) for a finite population or universe whose true value is unknown and $\hat{\theta}$ is some estimator (based on a probability sample, a census, a statistical model, an experiment, a scientific investigation, records, or data from various sources) whose value is known, the hope is that the difference $|\hat{\theta}-\theta|$ between $\hat{\theta}$ and θ is small. Assessment of an estimate's quality, possibly including sampling frame completeness, can include a study of the extent of such a difference, even though θ's true value is unknown. Such assessments can reveal limitations of the estimate, identify stages of the data collection process or data analysis that need improvement, and may lead to revelations of inadequacies in the frame (e.g., undercoverage, overcoverage, duplication, imperfect frame data, etc.) or differences in measurement concepts and definitions.

As Kish and Hess (1958) note, "The estimation of errors (differences) ... generally entails one of two difficult alternatives. The first calls for a quality check by procedures which are sufficiently better to provide the *true value* against which the (sample) survey results can be compared ... The alternative procedure calls for a reliable (and comparable) estimate from an outside source."

Disclaimer & Acknowledgments: The views expressed are those of the author and not necessarily those of the U. S. Bureau of the Census. The author is Chief, Center for Statistical Research & Methodology, U. S. Census Bureau and adjunct faculty in the Mathematics & Statistics Department, Georgetown University. The author is grateful to the following for technical reviews of earlier drafts: Patrick Cantwell, Jason Devine, Deborah Griffin, Heather King, Mary Mulry, and Victoria Velkoff. Special thanks to Alisha Armas and Michael Leibert for excellent research assistance.

From Tsao and Wright (1983), "... Morgenstern (1963) notes, there are several occasions when ... 'one is faced with alternative sets of data (estimates) which aim to describe the same phenomenon but which appear quite different; and it is difficult, if not impossible, to discover just where their difference lies.' Martin (1981) has pointed out how the comparisons of different data series produced by federal statistics bureaus can make important contributions to the improvements in one or both data series and that precautions are needed to avoid judgments based solely on supposition rather than on scientific evidence. Also, Eckler and Pritzker (1951) consider such external data comparisons as the first step in checking the adequacy of the frame."

Because there will usually be differences between definitions, target populations, responding populations, data collection procedures, and estimation procedures, the identification of differences in the estimates does not necessarily suggest that a particular estimate is inaccurate or that the other estimates are inaccurate. The value of a comparison lies in its ability to (1) give notice of changes in the population's structure occurring over time that may be affecting the estimates, (2) allow comparisons between different data collection procedures, and perhaps most importantly, (3) alert a particular agency to the need for further investigations.

1.2 Example Based on Counts and Estimates Associated with the 2010 Census

Let θ be the true total U. S. resident population as of April 1, 2010. The U. S. Census Bureau produced an official count $\hat{\theta}_1$ and three other estimates of θ as follows. Table 1 presents these four quantities, each aiming to estimate θ as close as possible.

During 2010, a nationwide constitutionally required census was conducted to distribute the 435 seats in the U. S. House of Representatives among the states. The table in the Appendix gives details of the construction of the 2010 Census count in terms of the various enumeration methods, which are defined following that table. During the 2010 Census, 300,758,215 persons were enumerated in housing units and 7,987,323 persons were enumerated in group quarters. The official 2010 Census resident population count delivered to the President of the United States on December 21, 2010 was $\hat{\theta}_1 = 308,745,538$.

The second estimate in Table 1 results primarily from an analysis of administrative records (e.g., births, deaths, and migration) often presented at a very high level as

$$\begin{aligned} \text{Population} &= \{\text{Population Under Age 65 in 2010}\} + \{\text{Population Ages 65 and Over in 2010}\} \\ &= \{\text{B - D + I - E}\} + \{\text{Medicare-based Population}\} \end{aligned}$$

where B = number of births since 1945; D = number of deaths of persons born since 1945; I = immigration of persons born since 1945; and E = emigration of persons born since 1945. This demographic analysis (DA) estimate was released by demographers on December 6, 2010 with value $\hat{\theta}_2 = 308,475,000$. The demographers actually released a spectrum of DA estimates with varying underlying migration assumptions:

305,684,000 • 307,415,000 • 308,475,000 • 310,038,000 • 312,713,000 .

The DA estimates were developed without knowledge of the 2010 Census counts.

The third estimate in Table 1 comes from the Census Bureau's Population Estimates Program (PEP). This estimate for 2010 is derived by starting with the Census 2000 count and updating it annually from 2000 to 2010 using administrative records of births, deaths, and migration. This third estimate, created without knowledge of the 2010 Census counts, was released by demographers in February 2011 with value $\hat{\theta}_3 = 308,450,484$.

The fourth estimate results from statistical methods that combine results from a nationwide probability sample [Census Coverage Measurement (CCM)] with the count from the 2010 Census, and it provides a measure of the quality of the 2010 Census count. This estimate was constructed from estimates released on May 22, 2012 with value $\hat{\theta}_4$ = 308,709,100 (= 300,667,000 estimated in housing units + 7,987,323 counted in group quarters + 54,777 counted in the remote Alaska area of which 54,111 were counted by the remote Alaska enumeration method as noted in the table given in the Appendix).

Table 1. Four Estimates of θ = the True Total U. S. Resident Population as of April 1, 2010.

Source	**Count or Estimate**
1. Official 2010 Census Resident Population Count	$\hat{\theta}_1$ = 308,745,538
2. Demographic Analysis (DA) Estimate	$\hat{\theta}_2$ = 308,475,000
3. Population Estimates Program (PEP) Estimate	$\hat{\theta}_3$ = 308,450,484
4. Probability Sample [Census Coverage Measurement (CCM)] Estimate	$\hat{\theta}_4$ = 308,709,100

Table 2 compares state counts from the 2010 Census with state 2010 PEP estimates. As illustrated in Table 2, analysts, when making initial comparisons, often will simply display competing estimates side-by-side and may compute the relative difference between them. Such comparisons are useful, but they are limited because: (1) they only show that two or more estimates differ, and (2) they only show the magnitude of the difference relative to a particular estimate.

Do these analyses directly address the issue of accuracy? These comparisons make statements about how close the estimates are to each other, but rarely, if ever, is any attempt made to say how close the set of estimates is to the true unknown value of the parameter of interest θ, which is what we ultimately want to know.

2. THE MAXIMUM RATIO

2.1. Concepts, Definitions, and Main Result

There is a simple and potentially useful tool called the maximum ratio which also gives an indication of how the collection of estimates stands relative to the unknown true value. That is, it provides some information about the accuracy of the collection of competing estimates. What follows provides highlights to the maximum ratio (Tsao and Wright, 1981, 1983).

Given the positive estimates $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m$ of the unknown positive value of the target parameter θ, we seek to focus on two questions: (1) How does one measure the closeness among the estimates? (2) Is at least one of the estimates unacceptable as an estimate of θ?

Alternatively to Table 2, ratios might form the basis of the measure of closeness. So

(1) if $\hat{\theta}_1$ is a statistical agency's estimate of primary interest, and if $\hat{\theta}_2, \hat{\theta}_3, ..., \hat{\theta}_m$ are competing estimates, then

(2) one could compute the ratios $\dfrac{\hat{\theta}_2}{\hat{\theta}_1}, \dfrac{\hat{\theta}_3}{\hat{\theta}_1}, ..., \dfrac{\hat{\theta}_m}{\hat{\theta}_1}$.

Steps (1) and (2) are appealing because they are simple and easy to explain.

Table 2. Comparisons between Preliminary Population Estimates and 2010 Census Counts for the States as of April 1, 2010. [*Source:* Population Division (February 2011), U.S. Bureau of the Census, Washington, D.C.]

Geographic Area	2010 Census Count	2010 PEP Estimate	2010 Census Count - 2010 PEP Estimate	(2010 Census Count − 2010 PEP Estimate) / 2010 Census Count × 100%
Alabama	4779736	4724112	55624	1.2
Alaska	710231	705175	5056	0.7
Arizona	6392017	6654358	-262341	-4.1
Arkansas	2915918	2904540	11378	0.4
California	37253956	37171135	82821	0.2
Colorado	5029196	5075295	-46099	-0.9
Connecticut	3574097	3523925	50172	1.4
Delaware	897934	889722	8212	0.9
District of Columbia	601723	607918	-6195	-1.0
Florida	18801310	18636368	164942	0.9
Georgia	9687653	9884534	-196881	-2.0
Hawaii	1360301	1296885	63416	4.7
Idaho	1567582	1555957	11625	0.7
Illinois	12830632	12931584	-100952	-0.8
Indiana	6483802	6438366	45436	0.7
Iowa	3046355	3019493	26862	0.9
Kansas	2853118	2835125	17993	0.6
Kentucky	4339367	4332584	6783	0.2
Louisiana	4533372	4519356	14016	0.3
Maine	1328361	1313697	14664	1.1
Maryland	5773552	5724856	48696	0.8
Massachusetts	6547629	6621588	-73959	-1.1
Michigan	9883640	9936913	-53273	-0.5
Minnesota	5303925	5283424	20501	0.4
Mississippi	2967297	2957749	9548	0.3
Missouri	5988927	6004372	-15445	-0.3
Montana	989415	978649	10766	1.1
Nebraska	1826341	1807012	19329	1.1
Nevada	2700551	2650677	49874	1.8
New Hampshire	1316470	1323202	-6732	-0.5
New Jersey	8791894	8723152	68742	0.8
New Mexico	2059179	2027191	31988	1.6
New York	19378102	19564202	-186100	-1.0
North Carolina	9535483	9432921	102562	1.1
North Dakota	672591	651787	20804	3.1
Ohio	11536504	11532245	4259	0.0
Oklahoma	3751351	3716212	35139	0.9
Oregon	3831074	3847469	-16395	-0.4
Pennsylvania	12702379	12625433	76946	0.6
Rhode Island	1052567	1056987	-4420	-0.4
South Carolina	4625364	4586078	39286	0.8
South Dakota	814180	817760	-3580	-0.4
Tennessee	6346105	6326403	19702	0.3
Texas	25145561	25101907	43654	0.2
Utah	2763885	2818242	-54357	-2.0
Vermont	625741	622191	3550	0.6
Virginia	8001024	7928720	72304	0.9
Washington	6724540	6727469	-2929	0.0
West Virginia	1852994	1824505	28489	1.5
Wisconsin	5686986	5664218	22768	0.4
Wyoming	563626	546821	16805	3.0
United States	308745538	308450484	295054	0.1

Now if $\dfrac{\hat{\theta}_j}{\hat{\theta}_1} \approx 1$ for $j = 2, 3, ..., m$, especially if the m estimates were obtained independently, the agency might infer that its estimate $\hat{\theta}_1$, or indeed all of the estimates are near θ. But there could be uniform bias as illustrated in Figure 1 where all estimates tend to overestimate the target θ.

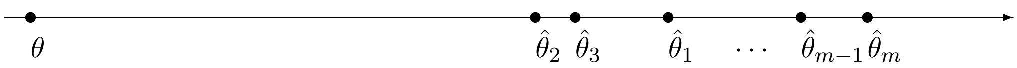
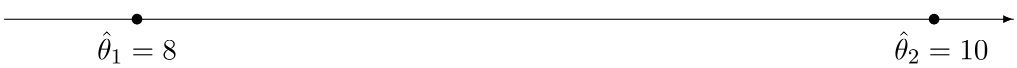


Figure 1. Uniform Positive Bias in All Estimates.

Such an approach, at best, implies a measure of closeness for the m estimates based on $\hat{\theta}_1$, but gives no information about the true value of θ.

Now consider a simple example of two competing estimates ($m = 2$) of a parameter θ:



Note that $\hat{\theta}_1 = 8$ is 20% of $\hat{\theta}_2 = 10$ away from $\hat{\theta}_2$; and that $\hat{\theta}_2 = 10$ is 25% of $\hat{\theta}_1 = 8$ away from $\hat{\theta}_1$.

If the desire is to report only one number as a measure of closeness for the two estimates, one natural (conservative) approach is to use the larger percentage because it gives an upper bound. Formally, we are led to the following definition (Tsao and Wright, 1981, 1983).

Definition 1: Let $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m$ be comparable estimates of the same parameter θ. The *maximum ratio* is given by:

$$M = \frac{\hat{\theta}_{(m)} - \hat{\theta}_{(1)}}{\hat{\theta}_{(1)}}$$

where $\hat{\theta}_{(1)} = min\{\hat{\theta}_1, ..., \hat{\theta}_m\}$ and $\hat{\theta}_{(m)} = max\{\hat{\theta}_1, ..., \hat{\theta}_m\}$.

The maximum ratio is a conservative measure of closeness because we divide by $\hat{\theta}_{(1)}$. Clearly, small values of the maximum ratio M imply that the estimates are close to each other. However, a small value of M for $\hat{\theta}_1, ..., \hat{\theta}_m$ does not imply that $\hat{\theta}_1, ..., \hat{\theta}_m$ are each close to θ.

Definition 2: Let $\theta > 0$ be the parameter of interest and $0 < \alpha < 1$. An estimate $\hat{\theta}$ (> 0) of θ is considered *unacceptable for θ for α* if $\hat{\theta} < (1 - \alpha)\theta$ or $\hat{\theta} > (1 + \alpha)\theta$. Alternatively, an estimate $\hat{\theta}(> 0)$ of θ is considered unacceptable for θ for α if $|\hat{\theta} - \theta| > \alpha\theta$.

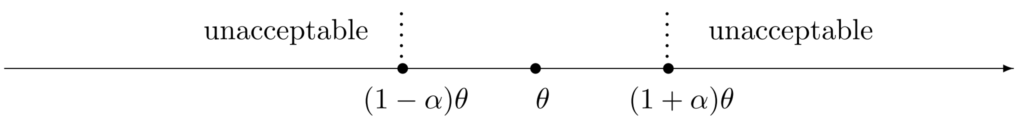


Figure 2. Unacceptable Regions for an Estimate $\hat{\theta}$.

It appears that in order to know if a particular estimate is unacceptable (i.e., too low or too high relative to θ) for a stated α, we must know the true value of θ, which we will never know. However, the presence of the competing estimates makes the following result possible. The proof is given in

Tsao and Wright (1983).

Main Result: If $M = M_0$ for m estimates of a particular parameter θ, then at least one of the estimates is unacceptable for θ for α if

$$\frac{M_0}{M_0+2} > \alpha \tag{1}$$

where $0 < \alpha < 1$ is some stated maximum tolerable error $\alpha\theta$.

While the main result is proved indirectly in Tsao and Wright (1981, 1983), this complete visual proof is motivated by a geometrical interpretation given in a personal communication Leff (1982).

Before the proof, we begin with some preliminaries. For any given set of estimates $\{\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m\}$, three possible relations exist between θ and the interval $(\hat{\theta}_{(1)}, \hat{\theta}_{(m)})$, and they are shown in Figure 3 (a, b, c). (Note that there are actually two versions of (a): the one given which shows θ closer to $\hat{\theta}_{(m)}$ and the one which is not shown would show θ closer to $\hat{\theta}_{(1)}$. The argument for the case of (a) which is not shown is similar to the case which is shown.) The letter E represents the amount of error for the *worst-estimate* of θ.

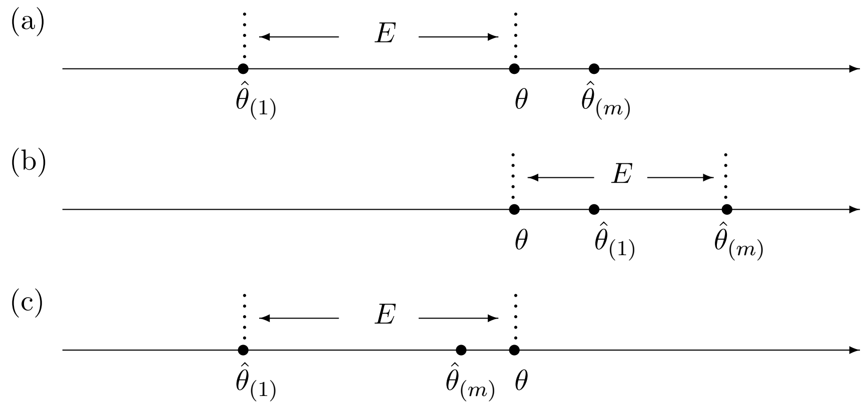


Figure 3. Three Possible Cases for the Worst-estimate $\hat{\theta}$ of θ.

Clearly from Figure 4 below, the *"best" worst-estimate* occurs when the interval $(\hat{\theta}_{(1)}, \hat{\theta}_{(m)})$ is such that θ is the midpoint of the interval. In this case, the error associated with $\hat{\theta}_{(1)}$ (and $\hat{\theta}_{(m)}$) is denoted by E^* and is called the *minimum worst-estimate error*. When θ is not the midpoint of the interval $(\hat{\theta}_{(1)}, \hat{\theta}_{(m)})$, then $E > E^*$.

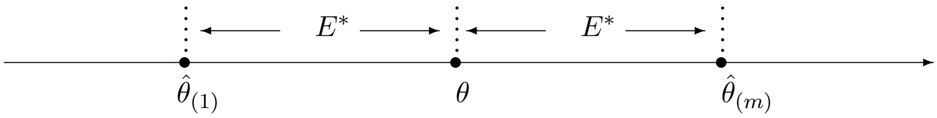


Figure 4. Illustration of "Best" Worst-estimate and Minimum Worst-estimate Error.

Now to give the proof of the *Main Result*, we consider two cases.

Case (i): When θ is the midpoint of the interval $(\hat{\theta}_{(1)}, \hat{\theta}_{(m)})$, then from Figure 4,

$$E^* = \theta - \hat{\theta}_{(1)} = \frac{M_0\hat{\theta}_{(1)}}{2} \tag{2}$$

which is equivalent to

$$M_0 = \frac{2(\theta - \hat{\theta}_{(1)})}{\hat{\theta}_{(1)}}. \tag{3}$$

If $\frac{M_0}{M_0+2} > \alpha$, then substituting (3) into $\frac{M_0}{(M_0+2)} > \alpha$ leads to $\hat{\theta}_{(1)} < (1-\alpha)\theta$ and hence $\hat{\theta}_{(1)}$ is unacceptable for θ for α.

Case (ii): When θ is *not* the midpoint of the interval $(\hat{\theta}_{(1)}, \hat{\theta}_{(m)})$, the error E of the worst-estimate is either $E = |\hat{\theta}_{(1)} - \theta|$ or $E = |\hat{\theta}_{(m)} - \theta|$ depending upon which case in Figure 3 holds. If $E = |\hat{\theta}_{(1)} - \theta|$ (similar for $E = |\hat{\theta}_{(m)} - \theta|$), we must have $\theta > \hat{\theta}_{(1)}$. So, because $E > E^*$ and using (2),

$$\theta - \hat{\theta}_{(1)} > \frac{\hat{\theta}_{(m)} - \hat{\theta}_{(1)}}{2} = \frac{M_o\hat{\theta}_{(1)}}{2} \tag{4}$$

which is equivalent to

$$M_o < \frac{2(\theta - \hat{\theta}_{(1)})}{\hat{\theta}_{(1)}}. \tag{5}$$

So if $\frac{M_o}{M_o+2} > \alpha$, then $M_o > \frac{2\alpha}{1-\alpha}$, which along with (5) leads to

$$\frac{2\alpha}{1-\alpha} < M_o < \frac{2(\theta - \hat{\theta}_{(1)})}{\hat{\theta}_{(1)}}. \tag{6}$$

Thus $\frac{\alpha}{1-\alpha} < \frac{\theta - \hat{\theta}_{(1)}}{\hat{\theta}_{(1)}}$ which implies $\hat{\theta}_{(1)} < (1-\alpha)\theta$. Hence $\hat{\theta}_{(1)}$ is unacceptable for θ for α.

Thus the *Main Result* is shown.

Comment: It should be noted that under *Case (i)* of the proof, the converse holds. That is, if $\hat{\theta}_{(1)}$ (similarly for $\hat{\theta}_{(m)}$) is unacceptable for θ for α, then from Definition 2,

$$\hat{\theta}_{(1)} < (1-\alpha)\theta$$

or equivalently

$$\theta - \hat{\theta}_{(1)} > \alpha\theta. \tag{7}$$

From (2)

$$\theta = \frac{M_0+2}{2}\hat{\theta}_{(1)},$$

and substituting this expression for θ into (7) gives (1), i.e.,

$$\frac{M_0}{M_0+2} > \alpha.$$

2.2. A Test of Unacceptability

The *Main Result* implies the following *test of unacceptability* for at least one of the given estimates.

Test of Unacceptability

---

1. State an α_0 that denotes a maximum tolerable error $\alpha_o\theta$ for the collection of estimates.

2. Obtain the primary estimate $\hat{\theta}_1$ and a set of competing estimates $\hat{\theta}_2, ..., \hat{\theta}_m$.

3. Compute the specific value of the maximum ratio $= M_0 = \dfrac{\hat{\theta}_{(m)} - \hat{\theta}_{(1)}}{\hat{\theta}_{(1)}}$.

4. Compare α_0 and $\dfrac{M_0}{M_0+2}$.

   (a) If $\dfrac{M_0}{M_0+2} > \alpha_0$, then at least one of the estimates is unacceptable for θ for α_0. Further investigation is needed.

   (b) If $\dfrac{M_0}{M_0+2} \leq \alpha_0$, then we lack sufficient evidence to say that at least one of the estimates is unacceptable for θ for α_0.

To illustrate the test of unacceptability, we return to the example of Section 1.2 where

$$\theta = \text{the true Total U. S. Resident Population as of April 1, 2010.}$$

1. Let's take $\alpha_0 = .005$. This means that we can not tolerate an estimate (or count) which is more than say $\frac{1}{2}$ of 1% of the true value of θ away from the true value of θ.

2. The U. S. Census Bureau produced an official count $\hat{\theta}_1$ and three other estimates as noted earlier in Table 1.

3. The maximum ratio is

$$M_0 = \frac{308,745,538 - 308,450,484}{308,450,484} = .000956568.$$

Thus the four estimates are within .096% of each other!

4. Finally, we have $\dfrac{M_0}{M_0+2} = .000478 \leq .005 = \alpha_0$. Hence we lack evidence to say that the official 2010 Census Count is unacceptable. The same holds for the other three estimates $\hat{\theta}_2, \hat{\theta}_3$, and $\hat{\theta}_4$. Thus with certainty, we lack evidence to declare that any of the four estimates is more than .5% of the true population away from the true population total.

Comment: Rather than state a pre-set maximum tolerable error $\alpha_0\theta$, one can just compute $\dfrac{M_0}{M_0+2}$. With this approach, any value of α less than $\dfrac{M_0}{M_0+2}$ would cause us to say at least one of the estimates is unacceptable for θ for α. For any value of α greater than or equal to $\dfrac{M_0}{M_0+2}$, we would lack sufficient evidence to make this claim.

Table 3 uses the test of unacceptability to compare state level estimates for the April 1, 2010 resident population. For each state, its population count from the 2010 Census is the primary estimate and a competing estimate comes from the Population Estimates Program (PEP).

Table 3. State Comparisons of PEP Estimates and 2010 Census Counts Using M_0

Geographic Area	2010 Census Count	2010 PEP Estimate	$M_0 \times 100\%$	$\frac{M_0}{M_0+2}$
Alabama	4779736	4724112	1.177449%	0.005853
Alaska	710231	705175	0.716985%	0.003572
Arizona	6392017	6654358	4.104197%	0.020108
Arkansas	2915918	2904540	0.391732%	0.001955
California	37253956	37171135	0.222810%	0.001113
Colorado	5029196	5075295	0.916628%	0.004562
Connecticut	3574097	3523925	1.423753%	0.007068
Delaware	897934	889722	0.922985%	0.004594
District of Columbia	601723	607918	1.029543%	0.005121
Florida	18801310	18636368	0.885054%	0.004406
Georgia	9687653	9884534	2.032288%	0.010059
Hawaii	1360301	1296885	4.889871%	0.023866
Idaho	1567582	1555957	0.747129%	0.003722
Illinois	12830632	12931584	0.786805%	0.003919
Indiana	6483802	6438366	0.705707%	0.003516
Iowa	3046355	3019493	0.889620%	0.004428
Kansas	2853118	2835125	0.634646%	0.003163
Kentucky	4339367	4332584	0.156558%	0.000782
Louisiana	4533372	4519356	0.310133%	0.001548
Maine	1328361	1313697	1.116239%	0.005550
Maryland	5773552	5724856	0.850607%	0.004235
Massachusetts	6547629	6621588	1.129554%	0.005616
Michigan	9883640	9936913	0.539002%	0.002688
Minnesota	5303925	5283424	0.388025%	0.001936
Mississippi	2967297	2957749	0.322813%	0.001611
Missouri	5988927	6004372	0.257893%	0.001288
Montana	989415	978649	1.100088%	0.005470
Nebraska	1826341	1807012	1.069666%	0.005320
Nevada	2700551	2650677	1.881557%	0.009320
New Hampshire	1316470	1323202	0.511368%	0.002550
New Jersey	8791894	8723152	0.788041%	0.003925
New Mexico	2059179	2027191	1.577947%	0.007828
New York	19378102	19564202	0.960362%	0.004779
North Carolina	9535483	9432921	1.087277%	0.005407
North Dakota	672591	651787	3.191840%	0.015709
Ohio	11536504	11532245	0.036931%	0.000185
Oklahoma	3751351	3716212	0.945560%	0.004706
Oregon	3831074	3847469	0.427948%	0.002135
Pennsylvania	12702379	12625433	0.609452%	0.003038
Rhode Island	1052567	1056987	0.419926%	0.002095
South Carolina	4625364	4586078	0.856636%	0.004265
South Dakota	814180	817760	0.439706%	0.002194
Tennessee	6346105	6326403	0.311425%	0.001555
Texas	25145561	25101907	0.173907%	0.000869
Utah	2763885	2818242	1.966688%	0.009738
Vermont	625741	622191	0.570564%	0.002845
Virginia	8001024	7928720	0.911925%	0.004539
Washington	6724540	6727469	0.043557%	0.000218
West Virginia	1852994	1824505	1.561465%	0.007747
Wisconsin	5686986	5664218	0.401962%	0.002006
Wyoming	563626	546821	3.073218%	0.015134
United States	308745538	308450484	0.095657%	0.000478

3. MAXIMUM RATIO: EXTENSION TO 2-D PARAMETERS

In this section, we extend the maximum ratio concept and *Main Result* to vector-valued parameters. As an example, consider the vector-valued parameter (θ, ϕ) where θ is the total value of sales by U.S. retail companies for month A and ϕ is the total value of inventories for U.S. retail companies for month A. Another example could be in terms of energy consumed per person living in housing units during month B (θ) and energy consumed per person living in group quarters during month B (ϕ).

Definition 3: The *error* in using $(\hat{\theta}, \hat{\phi})$ as an estimate of the vector-valued parameter (θ, ϕ) is the Euclidean distance between $(\hat{\theta}, \hat{\phi})$ and (θ, ϕ), i.e.,

$$Error = \sqrt{(\hat{\theta} - \theta)^2 + (\hat{\phi} - \phi)^2} \tag{8}$$

Definition 4: An estimate $(\hat{\theta}, \hat{\phi})$ of (θ, ϕ) is considered as *unacceptable* for (θ, ϕ) for α where $0 < \alpha < 1$, if the error in using $(\hat{\theta}, \hat{\phi})$ as an estimate of the truth (θ, ϕ) is at least α times the magnitude of (θ, ϕ), i.e., if

$$\sqrt{(\hat{\theta} - \theta)^2 + (\hat{\phi} - \phi)^2} > \alpha\sqrt{\theta^2 + \phi^2}. \tag{9}$$

Visually, or geometrically, assume a circle centered at (θ, ϕ) of radius $\alpha\sqrt{\theta^2 + \phi^2}$. If the estimate $(\hat{\theta}, \hat{\phi})$ is exterior to and beyond the circle, then $(\hat{\theta}, \hat{\phi})$ is unacceptable (Figure 5).

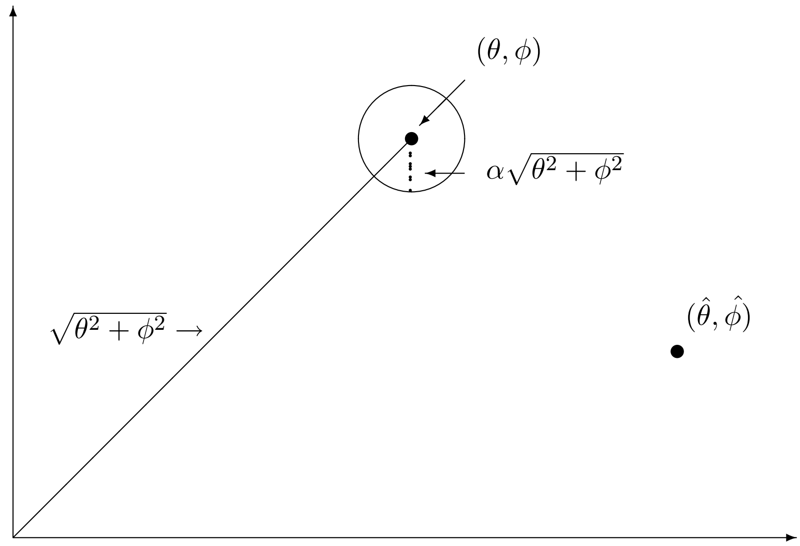


Figure 5. $(\hat{\theta}, \hat{\phi})$ is unacceptable for (θ, ϕ) for α.

Definition 5: Let $(\hat{\theta}_1, \hat{\phi}_1), (\hat{\theta}_2, \hat{\phi}_2), ..., (\hat{\theta}_m, \hat{\phi}_m)$ be m competing estimates of the same parameter (θ, ϕ) where $\theta > 0$ and $\phi > 0$. The *maximum ratio* that gives a measure of closeness among the competing estimates is given by

$$M = \frac{\sqrt{(\hat{\theta}_{(m)} - \hat{\theta}_{(1)})^2 + (\hat{\phi}_{(m)} - \hat{\phi}_{(1)})^2}}{\sqrt{\hat{\theta}^2_{(1)} + \hat{\phi}^2_{(1)}}} \tag{10}$$

where $\hat{\theta}_{(1)} = min\{\hat{\theta}_i\}$, $\hat{\phi}_{(1)} = min\{\hat{\phi}_i\}$, $\hat{\theta}_{(m)} = max\{\hat{\theta}_i\}$, and $\hat{\phi}_{(m)} = max\{\hat{\phi}_i\}$.

Example: Assume $(\theta, \phi) = (7, 9)$ and that three competing estimates are $(\hat{\theta}_1, \hat{\phi}_1) = (6, 10)$; $(\hat{\theta}_2, \hat{\phi}_2) = (7, 10)$; and $(\hat{\theta}_3, \hat{\phi}_3) = (8, 9)$. See Figure 6.

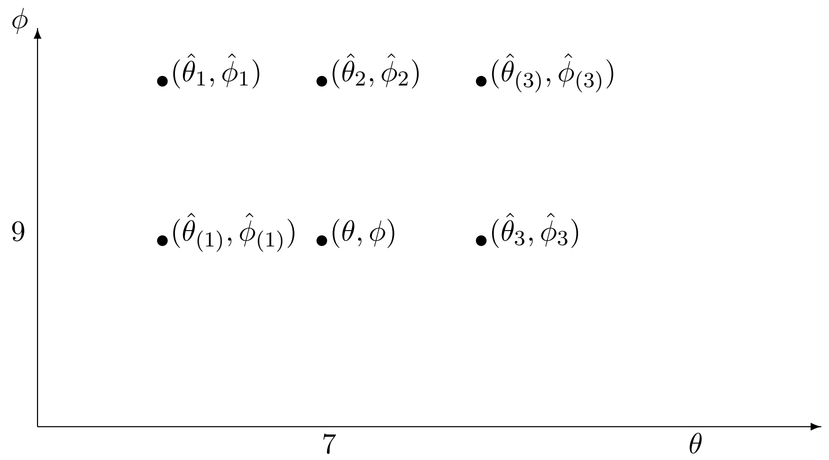


Figure 6.

Note that $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)}) = (6, 9)$ and $(\hat{\theta}_{(3)}, \hat{\phi}_{(3)}) = (8, 10)$. Furthermore, the distances between selected vector-valued estimates are as noted in Table 4:

Table 4. Distances Between Selected Vector-valued Estimates

Estimate 1	**Estimate 2**	**Distance**
$(\hat{\theta}_1, \hat{\phi}_1)$	$(\hat{\theta}_2, \hat{\phi}_2)$	$\sqrt{(6-7)^2 + (10-10)^2} = 1.$
$(\hat{\theta}_1, \hat{\phi}_1)$	$(\hat{\theta}_3, \hat{\phi}_3)$	$\sqrt{(6-8)^2 + (10-9)^2} = \sqrt{5}.$
$(\hat{\theta}_2, \hat{\phi}_2)$	$(\hat{\theta}_3, \hat{\phi}_3)$	$\sqrt{(7-8)^2 + (10-9)^2} = \sqrt{2}.$
$(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$	$(\hat{\theta}_{(3)}, \hat{\phi}_{(3)})$	$\sqrt{(6-8)^2 + (9-10)^2} = \sqrt{5}.$

Also the magnitudes of various vectors are

Table 5. Magnitudes of Selected Vector-valued Estimates

Vector	**Magnitude**
$(\hat{\theta}_1, \hat{\phi}_1)$	$\sqrt{6^2+10^2}=\sqrt{136}$
$(\hat{\theta}_2, \hat{\phi}_2)$	$\sqrt{7^2+10^2}=\sqrt{149}$
$(\hat{\theta}_3, \hat{\phi}_3)$	$\sqrt{8^2+9^2}=\sqrt{145}$
$(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$	$\sqrt{6^2+9^2}=\sqrt{117}$

Now relative to the magnitude of $(\hat{\theta}_1, \hat{\phi}_1)$, the points $(\hat{\theta}_1, \hat{\phi}_1)$ and $(\hat{\theta}_2, \hat{\phi}_2)$ are within $\frac{1}{\sqrt{136}} \times 100\% =$ 8.57% of each other. Similarly, we have

Table 6. Relative Closeness of Selected Vector-valued Estimates

Relative to ...	points ...			are within ... % of each other.
$(\hat{\theta}_2, \hat{\phi}_2)$	$(\hat{\theta}_1, \hat{\phi}_1)$	and	$(\hat{\theta}_2, \hat{\phi}_2)$	8.19%
$(\hat{\theta}_1, \hat{\phi}_1)$	$(\hat{\theta}_1, \hat{\phi}_1)$	and	$(\hat{\theta}_3, \hat{\phi}_3)$	19.17%
$(\hat{\theta}_3, \hat{\phi}_3)$	$(\hat{\theta}_1, \hat{\phi}_1)$	and	$(\hat{\theta}_3, \hat{\phi}_3)$	18.57%
$(\hat{\theta}_2, \hat{\phi}_2)$	$(\hat{\theta}_2, \hat{\phi}_2)$	and	$(\hat{\theta}_3, \hat{\phi}_3)$	11.59%
$(\hat{\theta}_3, \hat{\phi}_3)$	$(\hat{\theta}_2, \hat{\phi}_2)$	and	$(\hat{\theta}_3, \hat{\phi}_3)$	11.74%

However and relative to the magnitude of $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$, the points $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$ and $(\hat{\theta}_{(3)}, \hat{\phi}_{(3)})$ are within $\frac{\sqrt{5}}{\sqrt{117}} \times 100\% = 20.67\%$ of each other. Note that

$$M_0 \times 100\% = \frac{\sqrt{5}}{\sqrt{117}} \times 100\% = 20.67\%,$$

so the maximum ratio is thus a "conservative" measure of closeness among the three points $(\hat{\theta}_1, \hat{\phi}_1)$, $(\hat{\theta}_2, \hat{\phi}_2)$, and $(\hat{\theta}_3, \hat{\phi}_3)$. Thus we would say that the three estimates are within $M_0 \times 100\% = 20.67\%$ of each other.

For a given set of estimates $\{(\hat{\theta}_1, \hat{\phi}_1), ..., (\hat{\theta}_m, \hat{\phi}_m)\}$, consider the rectangle with vertices $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$, $(\hat{\theta}_{(1)}, \hat{\phi}_{(m)})$, $(\hat{\theta}_{(m)}, \hat{\phi}_{(1)})$, and $(\hat{\theta}_{(m)}, \hat{\phi}_{(m)})$ as shown in the Figure 7 below with two possibilities for (θ, ϕ). In case (a), (θ, ϕ) is inside the rectangle; and in case (b), (θ, ϕ) is outside the rectangle.

Definition 6: In either case (a) or (b) of Figure 7, E represents the *amount of error for the worst-estimate* of (θ, ϕ).

Definition 7: Clearly from Figure 8 below, the *"best" worst-estimate* occurs when $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$ and $(\hat{\theta}_{(m)}, \hat{\phi}_{(m)})$ are such that (θ, ϕ) is the midpoint of the rectangle with vertices as noted below. In this case, the error associated with $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$ [also the error for each of the other three vertices] is denoted by E^* and is called the *minimum worst-estimate error*.

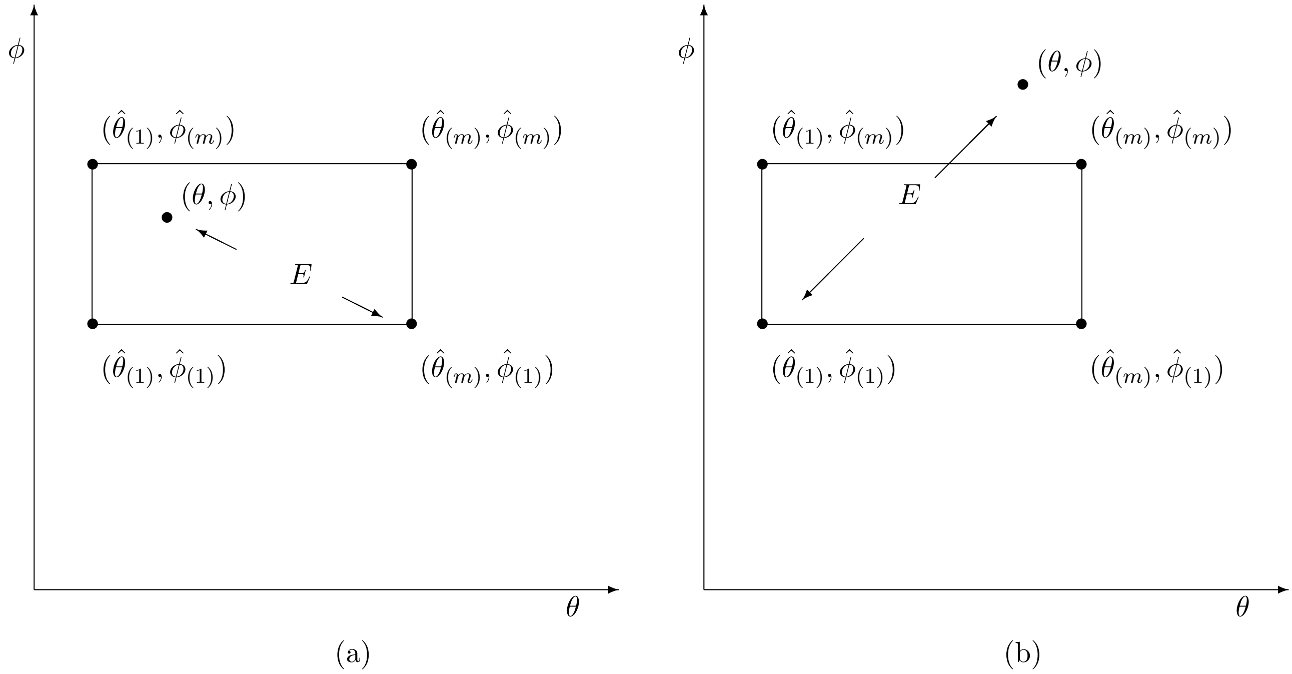


Figure 7. Two Possibilities for (θ, ϕ).

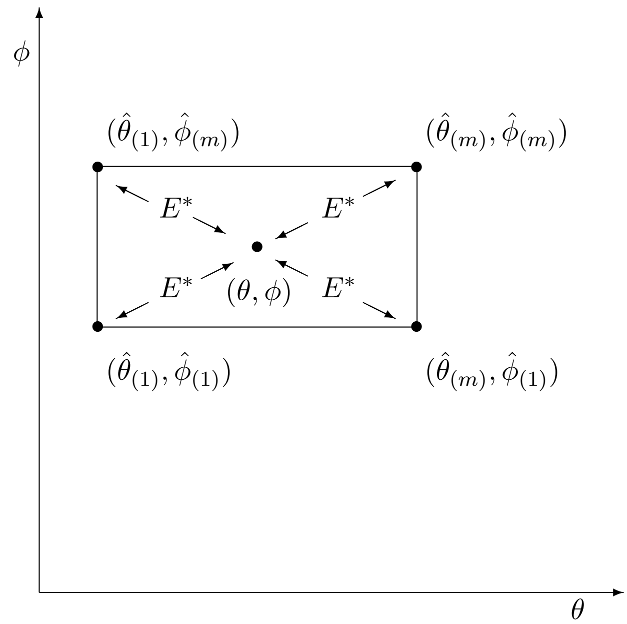


Figure 8. E^*, the Minimum Worst-estimate Error.

Main Result (for 2-D): If M has the particular value M_0 for m estimates $(\hat{\theta}_1, \hat{\phi}_1), (\hat{\theta}_2, \hat{\phi}_2), ..., (\hat{\theta}_m, \hat{\phi}_m)$ of a particular parameter (θ, ϕ), then at least one of the estimates $[(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$; $(\hat{\theta}_{(1)}, \hat{\phi}_{(m)})$; $(\hat{\theta}_{(m)}, \hat{\phi}_{(1)})$; $(\hat{\theta}_{(m)}, \hat{\phi}_{(m)})]$ is unacceptable [i.e., more than $100\alpha\%$ of the magnitude of (θ, ϕ) away from (θ, ϕ)] for α if

$$M_0\sqrt{\frac{\hat{\theta}^2_{(1)} + \hat{\phi}^2_{(1)}}{(\hat{\theta}_{(1)} + \hat{\theta}_{(m)})^2 + (\hat{\phi}_{(1)} + \hat{\phi}_{(m)})^2}} > \alpha. \tag{11}$$

Proof: Now to give the proof of the *Main Result (for 2-D)*, we consider two cases.

Case (i). When (θ, ϕ) is the midpoint of the 'rectangle' as given in Figure 8:

For the case in Figure 8 and focusing on $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$, similarly for the other three vertices, we have the minimum worst estimate error is

$$E^* = \sqrt{(\hat{\theta}_{(1)} - \theta)^2 + (\hat{\phi}_{(1)} - \phi)^2} = \frac{M_0\sqrt{\hat{\theta}^2_{(1)} + \hat{\phi}^2_{(1)}}}{2}. \tag{12}$$

That is, E^*, the minimum worst estimate error, equals half the length of the diagonal.

We also see from Figure 8 that

$$(\theta, \phi) = (\frac{\hat{\theta}_{(1)} + \hat{\theta}_{(m)}}{2}, \frac{\hat{\phi}_{(1)}, \hat{\phi}_{(m)}}{2}) \tag{13}$$

because (θ, ϕ) is the midpoint between the points $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$ and $(\hat{\theta}_{(m)}, \hat{\phi}_{(m)})$.

Thus if

$$M_0\sqrt{\frac{\hat{\theta}^2_{(1)} + \hat{\phi}^2_{(1)}}{(\hat{\theta}_{(1)} + \hat{\theta}_{(m)})^2 + (\hat{\phi}_{(1)} + \hat{\phi}_{(m)})^2}} > \alpha,$$

then

$$M_o\sqrt{\theta^2_{(1)} + \phi^2_{(1)}} > \alpha\sqrt{(\hat{\theta}_{(1)} + \hat{\theta}_{(m)})^2 + (\hat{\phi}_{(1)} + \hat{\phi}_{(m)})^2}$$

which becomes

$$M_o\sqrt{\theta^2_{(1)} + \phi^2_{(1)}} > 2\alpha\sqrt{(\frac{\hat{\theta}_{(1)} + \hat{\theta}_{(m)}}{2})^2 + (\frac{\hat{\phi}_{(1)} + \hat{\phi}_{(m)}}{2})^2}. \tag{14}$$

Solving for M_o in (12) and substituting the value for M_o into (14) and substituting the result from (13) gives

$$(\frac{2\sqrt{(\hat{\theta}_{(1)} - \theta)^2 + (\hat{\phi}_{(1)} - \phi)^2}}{\sqrt{\hat{\theta}^2_{(1)} + \hat{\phi}^2_{(1)}}})\sqrt{\theta^2_{(1)} + \phi^2_{(1)}} > 2\alpha\sqrt{\theta^2 + \phi^2}$$

which reduces to

$$\sqrt{(\hat{\theta}_{(1)} - \theta)^2 + (\hat{\phi}_{(1)} - \phi)^2} > \alpha\sqrt{\theta^2 + \phi^2}.$$

Hence $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$ is unacceptable for (θ, ϕ) for α.

Conversely, if $(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$ is unacceptable for (θ, ϕ) for α, then

$$\sqrt{(\hat{\theta}_{(1)} - \theta)^2 + (\hat{\phi}_{(1)} - \phi)^2} > \alpha\sqrt{\theta^2 + \phi^2}$$

which is equivalent to

$$\left(\frac{2\sqrt{(\hat{\theta}_{(1)} - \theta)^2 + (\hat{\phi}_{(1)} - \phi)^2}}{\sqrt{\hat{\theta}^2_{(1)} + \hat{\phi}^2_{(1)}}}\right)\sqrt{\theta^2_{(1)} + \phi^2_{(1)}} > 2\alpha\sqrt{\theta^2 + \phi^2} \tag{15}$$

Using M_o from (12) and the result in (13), (15) becomes

$$M_o\sqrt{\theta^2_{(1)} + \phi^2_{(1)}} > \alpha\sqrt{(\hat{\theta}_{(1)} + \hat{\theta}_{(m)})^2 + (\hat{\phi}_{(1)} + \hat{\phi}_{(m)})^2}$$

which is equivalent to

$$M_0\sqrt{\frac{\hat{\theta}^2_{(1)} + \hat{\phi}^2_{(1)}}{(\hat{\theta}_{(1)} + \hat{\theta}_{(m)})^2 + (\hat{\phi}_{(1)} + \hat{\phi}_{(m)})^2}} > \alpha. \tag{16}$$

Thus we have shown that in *Case (i)* when (θ, ϕ) is the midpoint of the rectangle, then the minimum worst estimate error E^* associated with a 'best' worst estimate is unacceptable if and only if (11) holds.

Case (ii). When (θ, ϕ) is 'not' the midpoint of the 'rectangle' as given in Figure 8:

If (θ, ϕ) is not the midpoint of the rectangle, then the error E of a worst-estimate is one of the following:

$$E = \sqrt{(\hat{\theta}_{(1)} - \theta)^2 + (\hat{\phi}_{(1)} - \phi)^2};$$
$$E = \sqrt{(\hat{\theta}_{(1)} - \theta)^2 + (\hat{\phi}_{(m)} - \phi)^2};$$
$$E = \sqrt{(\hat{\theta}_{(m)} - \theta)^2 + (\hat{\phi}_{(1)} - \phi)^2}; \text{ or}$$
$$E = \sqrt{(\hat{\theta}_{(m)} - \theta)^2 + (\hat{\phi}_{(m)} - \phi)^2}.$$

Whatever the case, $E \geq E^*$.

Now we have just seen that E^* is the minimum worst error associated with the 'best' worst-estimate, and it is unacceptable if and only if (11) holds. So if (11) holds, then E^* is unacceptable error and because $E \geq E^*$, this implies that E is also unacceptable error for at least one of the vertices of the rectangle for α. Thus at least one of the following is unacceptable for (θ, ϕ) for α:

$(\hat{\theta}_{(1)}, \hat{\phi}_{(1)})$; $(\hat{\theta}_{(1)}, \hat{\phi}_{(m)})$; $(\hat{\theta}_{(m)}, \hat{\phi}_{(1)})$; or $(\hat{\theta}_{(m)}, \hat{\phi}_{(m)})$.

Thus the *Main Result (for 2-D)* is shown.

4. CONCLUDING REMARKS

4.1. A Standard for Census Assessment: The *maximum ratio* and the *test of unacceptability* hold promise as useful tools for developing a *standard* against which to assess, at a very high level, the success or quality of a decennial census relative to the unknown truth when α is appropriately small. Given $\frac{M_0}{M_0+2}$ for $\hat{\theta}_1$ the official decennial census count and some competing estimates $\hat{\theta}_2, \hat{\theta}_3, ..., \hat{\theta}_m$, for any value of α_0 where $\frac{M_0}{M_0+2} \leq \alpha_0$, the Census Bureau would lack evidence to say that $\hat{\theta}_1$ is unacceptable. For any α_0 where $\frac{M_0}{M_0+2} > \alpha_0$, at least one of the estimates in the collection $\{\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m\}$ is unacceptable, and hence further investigation is required.

4.2. Other Official Statistics Applications: The test of unacceptability can be applied in any situation where there is a collection of competing estimates for the same target parameter. For official statistics, many potential applications exist. For example, the decennial census provides a national vacancy rate of housing units which can be compared with similar estimates from sample surveys. Estimates of energy sales volumes could be compared with estimates of energy consumption and storage volumes. Estimates of purchases by consumers could be compared with estimates of sales by businesses. Estimates of food stamp purchases from household sample surveys could be compared to estimates of food stamp funds distributed by the responsible government agency.

4.3. Reproducibility in Science: Sometimes, scientific measurement yields many widely ranging estimates. One recent case involves global warming. With estimates of the rate at which ice sheets in Antartica and Greenland have been losing (or gaining) mass ranging from an overall loss of 676 billion tons per year to a gain of 69 billion tons a year, 47 experts from 26 institutions around the world reviewed data from over 25 studies to arrive at a reconciled estimate (Kerr, 2012; Shepherd, et al, 2012) by considering the varying conditions under which the different estimates were obtained. In this case of global warming, the need for further investigation is clear.

More generally, the test of unacceptability can be applied in applications regarding questions of reproducibility of scientific results. A quick scan of some recent issues of the journal *Science* reveals several examples: (1) modern estimates of the number of eukaryotic (non-bacterial) species on Earth, ranging from 2 million to 100 million (Costello, et al, 2013); (2) different estimates of the root-mean-square proton charge radius from three methods (Antognini, et al, 2013; and Margolis, 2013); and (3) reference to a range of estimates between 6 and 37 petabecquerels of cesium-137 that was spewed by the March 2011 Fukushima (Japan) nuclear disaster (Normile, 2013).

The results on measuring the size of a proton, actually considering three different radii of the proton (the Zemach radius, the magnetic radius, and the charge radius), were investigated. We focus on charge radius (r_E) which is the root mean square charge radius given in femtometers (fm). Three previous estimates of the true unknown value of the charge radius of a proton r_E are the first three estimates given in Table 7.

Table 7. Four Estimates of the Charge Radius r_E of a Proton

Method	Estimates of r_E
1. Electron-Proton Scattering	0.879 fm
2. Electron-Proton Scattering	0.875 fm
3. Hydrogen Spectroscopy CODATA (2010)	0.8775 fm $\approx$ 0.878 fm
4. Muonic Hydrogen Spectroscopy	0.84087 fm $\approx$ 0.841 fm

The fourth estimate in Table 7 is obtained from new methodology and is reported in Antognini, et al (2013). Note that $M_0 = \frac{max\{r_E\} - min\{r_E\}}{min\{r_E\}} = \frac{.879 - .841}{.841} = .0452$. Thus the four estimates are within 4.52% of each other. Also $\frac{M_0}{M_0+2} = .0221$. So for any $\alpha_0 < .0221$, at least one of the four estimates is unacceptable and further investigation is required. While the fourth estimate seems far away from the other three, it may very well be the best of the four due to improved methodology. While noting that their new estimate of 0.84087 fm is in agreement with another estimate of $r_E = 0.84$ fm which results from "...a global fit of proton and neutron form factors based on dispersion relations and the vector-dominance model," Antognini, et al (2013) and Margolis(2013) call for further investigations to further explain the differences. [The Zemach radius (r_Z) has estimates of 1.086 fm, 1.045 fm, 1.047 fm, 1.037 fm, and 1.082 fm. Thus $M_0 = 0.0473$ and $\frac{M_0}{M_0+2} = 0.0231$. On the other hand, the magnetic radius (r_M) has estimates of $0.803 \approx 0.80$ fm, $0.867 \approx 0.87$ fm, 0.86 fm, and 0.87 fm. Thus $M_0 = 0.0875$ and $\frac{M_0}{M_0+2} = 0.0419$.]

4.4. A Deterministic Test: Finally, the test of unacceptability is deterministic and not stochastic. However, setting a value of α_0 is similar to setting an α-level of significance as in statistical hypothesis testing. Because the test of unacceptability is not stochastic, no adjustments are needed when repeating the test of unacceptability to subsets of estimates when determining which groupings produce unacceptable worst estimate error and which groupings do not. The maximum ratio and the test of unacceptability are particularly useful when the uncertainty of some of the considered estimates can not be quantified.

5. REFERENCES

- Antognini, A., et al. (2013). "Proton Structure from the Measurement of 2S-2P Transition Frequencies of Muonic Hydrogen," *Science, 339*, 417-420.
- Cantwell, P. (2012). "2010 Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Persons in the United States," *DSSD 2010 Census Coverage Measurement Memorandum Series # 2010-G-01*, U. S. Census Bureau, Washington, D.C.
- *Census Bureau Releases Estimates of Undercount and Overcount in the 2010 Census (5/22/2012)*, U. S. Census Bureau Press Release, Washington, D.C.
- Costello, M., May, R., and Stork, N. (2013). "Can We Name Earth's Species Before They Go Extinct?," *Science, 339*, 413-416.
- Eckler, A. R. and Pritzker, L. (1951). "Measuring the Accuracy of Enumerative Surveys," *Bulletin of the International Statistical Institute, Part 4, 33*, 7-24.
- Jackson, G. (2013). "The Number of Housing Units and People Counted in the 2010 Census by Operation," *DSSD 2020 Decennial Census R&T Memorandum Series # R-01* , U. S. Census Bureau, Washington, D.C.
- Kerr, R. (2012). "Experts Agree Global Warming Is Melting the World Rapidly," *Science, 338*, 1138.
- Kish, L. and Hess, I. (1958). "On Noncoverage of Sample Dwellings," *Journal of the American Statistical Association, 53*, 509-524.
- Leff, H. S. (1982). "Personal Correspondence," Institute for Energy Analysis, Oak Ridge Associated Universities, Oak Ridge, TN.
- Margolis, H. (2013). "How Big Is the Proton?," *Science, 339*, 405-406.

- Martin, M. E. (1981). "Statistical Practice in Bureaucracies," *Journal of the American Statistical Association, 76*, 1-8.
- Morgenstern, O. (1963). *On the Accuracy of Economic Observations*, Princeton, NJ: Princeton University Press.
- Mule, V. T. (2012). "U. S. Census Coverage Measurement Survey Results," *Proceedings of the Joint Statistical Meetings*, American Statistical Association, Alexandria, VA, 3425-3439.
- Normile, D. (2013). "Cooling a Hot Zone," *Science, 339*, 1028-1029.
- Shepherd, A., et al (2012). "A Reconciled Estimate of Ice-Sheet Mass Balance," *Science, 338*, 1183-1189.
- *The Development and Sensitivity Analysis of the 2010 Demographic Analysis Estimates (11/29/2010)*, Population Division, U. S. Census Bureau, Washington, D.C.
- Tsao, H. and Wright, T. (December 1981). "A Note on a Simple Tool for the Comparison of Several Estimates of the Same Parameter," *Oak Ridge National Laboratory/TM-7973*, Oak Ridge, TN.
- Tsao, H. and Wright, T. (1983). "On the Maximum Ratio: A Tool for Assisting Inaccuracy Assessment," *The American Statistician, 37:4*, 339-342.
- *U. S. Census Bureau Announces 2010 Census Population Counts; Apportionment Counts Delivered to President (12/21/2010)*, U. S. Census Bureau Press Release, Washington, D.C.
- Wright, T. and Cobb, G. (2005). "Counting and Apportionment: Foundations of America's Democracy," in *Statistics: A Guide to the Unknown (4^{th} ed.)*, eds. R. Peck, G. Casella, G. Cobb, R. Hoerl, D. Nolan, R. Starbuck, and H. Stern, Pacific Grove, CA: Duxbury Press, 35-67.

APPENDIX

Table. Official U.S. Resident Population Counted in the 2010 Census by Enumeration Method [*Source:* Jackson, 2013.]

Enumeration Method	**Number of Persons**
1. IN HOUSING UNITS	
Self-administered Questionnaires	
Mailout/Mailback	193,345,244
Update/Leave	13,610,355
Form Fulfillment	312,463
Be Counted Forms	303,041
Interviewer-administered Questionnaires	
Nonresponse Follow-up	66,785,373
Vacant Delete Check	4,665,512
Nonresponse Follow-up Residual	728,374
Update Enumerate	1,947,532
Remote Update Enumerate	5,967
Remote Alaska	54,112
Coverage Follow-up	17,326,236
Telephone Questionnaire Assistance	104,489
Enumeration of Transitory Locations	190,627
Nonresponse Follow-up Reinterview	157,410
Update Enumerate Reinterview	6,465
Imputation Cases and UHE Moves with an Unknown Operation Origin	1,215,015
2. IN GROUP QUARTERS	
Subtotal for Group Quarters	7,987,323
TOTAL RESIDENT POPULATION (April 1, 2010)	308,745,538

Glossary of 2010 Census Enumeration Methods (Jackson, 2013)

Mailout/Mailback: A 2010 Census questionnaire was mailed to a housing unit, and the housing unit completed the questionnaire and mailed it back.
Update/Leave: A 2010 Census field operation where a lister (of addresses) leaves a questionnaire at the housing unit for the people living at the housing unit to complete and mail back to the Census Bureau.
Form Fulfillment: A person calls Telephone Questionnaire Assistance and requests that a 2010 Census questionnaire be mailed to his or her address. The Form Fulfillment questionnaire was available in the following languages: English, Spanish, Chinese, Korean, Russian, and Vietnamese.
Be Counted: A person picked up, completed and sent back a 2010 Census Be Counted Questionnaire from a Questionnaire Assistance Center or Be Counted location. The Be Counted questionnaire was available in the following languages: English, Spanish, Chinese, Korean, Russian, and Vietnamese.
Nonresponse Follow-up: A person did not complete a mailback questionnaire and was interviewed by a Census enumerator.
Vacant Delete Check: A follow-up interview for housing units found to be non-seasonally vacant and nonexistent in Nonresponse Follow-up. Additionally, if a person mailed back a blank 2010 Census mailback questionnaire or if a housing unit was in the supplemental enumeration universe delivery, they were interviewed for the first time in the Vacant Delete Check operation. The supplemental enumeration universe includes housing units that were not known in time to be included in the initial enumeration universe. They include housing units that were added from operations like New Construction and Local Update of Census Address Appeals. Housing units in the supplemental universe were mailed questionnaires in the Late Add Mailout.
Nonresponse Follow-up Residual: If a Nonresponse Follow-up enumerator was unable to determine how many people lived in a housing unit, that housing unit was reinterviewed in the Nonresponse Follow-up Residual operation. Additionally, if a person mailed back a blank 2010 Census mailback questionnaire and was not included in the Vacant Delete Check operation, they were followed up with in the Nonresponse Follow-up Residual operation.
Update Enumerate: A Census Bureau enumerator canvassed and interviewed housing units located on

Native American reservations and areas that had a large amount of seasonally occupied housing units.
Remote Update Enumerate: A Census Bureau enumerator canvassed and interviewed housing units located in rural Maine and rural Alaska.
Remote Alaska: A Census Bureau enumerator canvassed and interviewed housing units located in extremely remote locations in Alaska. The operation was started in January and conducted in three waves.
Coverage Follow-up: A follow-up interview with respondents that were identified as having potential overcoverage or undercoverage on their mailback or Nonresponse Follow-up questionnaire.
Telephone Questionnaire Assistance: A person completed his or her 2010 Census questionnaire over the phone with an interviewer at a call center.
Enumeration of Transitory Locations: People who lived in transitory locations such as boat docks, hotels, and trailer parks and were interviewed by Census Bureau enumerators.
Reinterview: A quality control follow-up interview with respondents from Nonresponse Follow-up and Update Enumerate. The only reinterview questionnaires included in the final 2010 Census counts are those that replaced production questionnaires due to the production enumerator falsifying data or not following procedures correctly.
Group Quarters: A group quarters is a place where people live or stay that is normally owned or managed by an entity or organization providing housing or services for the residents. These services may include custodial or medical care as well as other types of assistance, and residency is commonly restricted to those receiving these services. People living in group quatrers are usually not related to one another. Group quarters include such places as college residence halls, residential treatment centers, skilled nursing facilities, group homes, military barracks, correctional facilities, workers' dormitories, and facilities for people experiencing homelessness. People in quarters are interviewed during the Group Quarters Enumeration operation.
UHE (Usual Household Elsewhere) Moves: A housing unit that is occupied by people that completed either a Be Counted Questionnaire or Group Quarters questionnaire and indicated that they stayed at another address.
Late Add Mailout: There were housing units identified too late to be included in the original 2010 Census mailout. These units were part of the supplemental enumeration universe. There was a separate mailout to these units, and they were assigned a Processing Identifier for the process instead of a Census Identifier. The Processing Identifier identifies housing units that were added to the initial enumeration universe.
Unlinked Continuation Forms: The 2010 Census questionnaires used by enumerators only include enough space to have information for up to five people living at the housing unit. If more than five people lived at a housing unit, a continuation form was used. The continuation forms were then linked to the parent form. However, not all continuation forms were able to be linked. An unlinked continuation form indicates that there was no parent form captured for this unit.
Imputation: The assignment of values by the Census Bureau when information is missing or inconsistent. Imputation relies on the tendency of households of the same size within a small geographic area to be similar in most characteristics.
Status Imputation: The Census Bureau did not know if the address was a housing unit or another structure. A value of existing or nonexistent was assigned to the housing unit.
Occupancy Imputation: The Census Bureau knew that the housing unit existed but its status was not known. A value of occupied or vacant was assigned to the housing unit.
Household Size Imputation: The Census Bureau knew that a housing unit was occupied but did not know how many people lived in the unit. The number of people living in that unit was assigned to it.