M^3 Challenge Third Place, Cum Laude Team Prize of $10,000

Maggie Walker Governor's School, Team #365
Richmond, VA

Coach: Kristine Chiodo

Students: Susan Margaret Ballentine, William Steadman Farmer, Ashish Ashok Makadia, Cody Tyler Talmadge, Milton Frederick Tyler IV
THE SENSIBLE CENSUS

SUMMARY

As mandated by our Constitution, the United States Government requires an "enumeration" every ten years to determine the number of people in the country so that it can apportion seats of the House of Representatives appropriately, a process we call the Census. Despite the best efforts of the U.S. Government, our nation's census remains inaccurate—the count is always below the true number, which is estimated by the Accuracy and Coverage Evaluation. To make up for those missed in the count, the government has imputed data; however, even this number is inaccurate. Our goal was to compare and devise a method to more accurately predict the actual population of the United States. We were also asked to find a method to fairly apportion the Congressional seats and find a method to redistrict each state while avoiding gerrymandering, which are both processes resulting from the census. In our approach, we looked at each section of the problem separately; although each previous problem impacts the next in a given situation, they can be looked at independently.

For our first model, we decided to use a mark and recapture method, used by the Census Bureau with the Accuracy and Coverage Evaluation (ACE), to readjust the census data. According to our calculations, using a mark and recapture method will yield less error than using the census alone to estimate the population. To test the model, we compared the standard error of the mark and recapture method to the undercount error in the 2000 Census, using data from the 2000 Census and ACE. From data given by the Statistical Research Center of the Census Bureau, the standard error of the ACE was found to be 541,631, about .19% of the population. This error is less than the number of people undercounted in the 2000 Census, 1,331,656 people, about .48% of the population.

We proposed that the best method for reapportionment utilizes the CAWES Model which is based on population size and rate of population increase. In most other apportionment models, seats are distributed evenly among all the states, giving each state a number exactly proportional to their populations. However, this number is a decimal, and because partial seats cannot be distributed, different models have been created to determine how these seats are given out. The CAWES model rounds this decimal number down and then gives seats back; due to the Constitutional requirement, it first gives those states with zero seats one. Then it distributes the leftover seats based on the population growth rates of each state. Our model contains many discrepancies with the currently used model (21 states differ in their apportionments). However, all differences between other models and the CAWES Model differ by only one seat, except Texas which gains two more with our method due to its increasing population.

It is proposed that the best method for redistricting the provincial cake between the Democratic and Republican Parties is the modified Zeph–Landau method. This model stems off the philosophy "I cut, you choose" in order to prevent the unfair distribution of districts and thus gerrymandering. The modified Zeph–Landau method allows states that are not equally party- distributed to have slanted districts, yet not let some districts consistently dominate others. This model could be applied in every state and is an ideal answer to stop gerrymandering.
# TABLE OF CONTENTS

SUMMARY ..................................................................................................... 1  
TABLE OF CONTENTS ............................................................................... 2  
LIST OF FIGURES ..................................................................................... 3  
INTRODUCTION .......................................................................................... 4  
  Background ............................................................................................... 4  
  Restatement ............................................................................................... 4  
  Global Assumptions .................................................................................. 4  
CENSUS UNDERCOUNTS .......................................................................... 4  
  Rationale .................................................................................................... 4  
  Assumptions ............................................................................................... 5  
  Design .......................................................................................................... 5  
  Error Analysis ........................................................................................... 6  
  Testing the Model....................................................................................... 6  
APPORTIONMENT ...................................................................................... 7  
  Rationale .................................................................................................... 7  
  Assumptions ............................................................................................... 7  
  Design .......................................................................................................... 7  
  Model .......................................................................................................... 8  
  Results ......................................................................................................... 9  
CONGRESSIONAL DISTRICTS .............................................................. 11  
  Rationale .................................................................................................... 11  
  Assumptions ............................................................................................... 11  
  Design .......................................................................................................... 11  
CONCLUSIONS ........................................................................................... 14  
REFERENCES ............................................................................................. 16
LIST OF FIGURES

Figure 1 - The CAWES Model with the predicted values for 2005, 2010, and 2015 9
Figure 2 - An example of a state being divided by population before redistricting occurs 12
Figure 3 - An example of two possible dividing lines 12
Figure 4 - An example of a possible distribution 13
Figure 5 - An example of several divisions and the resulting party win distribution 13
INTRODUCTION

Background

According to the Constitution of the United States of America, an "enumeration shall be made within three years after the first meeting of the Congress of the United States, and within every subsequent term of ten years, in such manner as they shall by law direct." A census accounts for the population from each district from each state as well as those individuals overseas given by employing federal departments and agencies from their administrative records.\(^1\) The census does not include private citizens that live abroad. Census presents "information about race, Hispanic origin, age, sex, household type, housing tenure, and other social, economic, and housing characteristics."

Restatement

This paper addresses the following questions: (1) How should the census figures be adjusted for the undercount of population? (2) What are the errors of our solution and how are they calculated? (3) What method should Congress select for apportioning the House of Representatives and how does this method compare to other methods? (4) What recommendations should be made to the states to ensure that Congressional districts are fairly drawn, and how are these recommendations justifiable?

Global Assumptions

1. There will be no loss or gain in states during the time the Census is taken.
2. There will be no sudden unexpected gains or losses in population.

CENSUS UNDERCOUNTS

Rationale

The Census Bureau’s estimate of the population of the United States for April 1, 2000 reported 278 million people living in the United States. The actual census count as announced by the Census Bureau on December 28, 2002 was 281.4 million people. Over three million people had not been accounted for in 2000.\(^2\) In the 2000 Census, "approximately 5.77 million persons had all their characteristics imputed" and were added to the census with imputations. Currently, "undercounts cost more than $26,000 [in lost federal funding] per 1000 people not counted."\(^3\)

---

2. [http://www.psc.isr.umich.edu/dis/acs/handouts/SupremeCourt.doc](http://www.psc.isr.umich.edu/dis/acs/handouts/SupremeCourt.doc)
3. [http://news.newamericamedia.org/news/view_article.html?article_id=e3183a93b659c0a19ebf4c439f0fe64b](http://news.newamericamedia.org/news/view_article.html?article_id=e3183a93b659c0a19ebf4c439f0fe64b)
There is much debate over imputations mainly because one party may be favored over another. "In general, poorer neighborhoods have higher undercount rates" because of "limited English proficiency,"^4^ and those individuals may sway toward a given party.

There are numerous reasons for the undercount in the 2000 Census, many of which will be repeated in the 2010 Census. One reason that people are missed in the census is because not everyone has a Decennial Master Address File (MAF) ID.^5^ Also, the form the census uses only has six people per household; if the household is larger, the extra people can only be accounted for by follow-up interviews and phone calls, which are unreliable. "Some immigrant workers have told ethnic newspapers and radio programs that they fear personal information could be used against them if it is revealed to local authorities—even if the information turns out to be inaccurate."^6^ Another problem is that if the method for collecting the census data is not completed by the deadline, people's information will be left out.

To prevent this major undercount, imputations should be taken into account. The U.S. Census Bureau created the Accuracy and Coverage Evaluation (ACE) in order to allow for the possibility of correcting the census results for the measured undercount. Currently there are three imputation types: The Household Size Imputations are used to estimate the number of persons in a household when occupancy is ambiguous. The Occupancy Imputations are used to estimate the population for housing units that do not have sufficient info to classify them as occupied or vacant. The Status Imputations are used to estimate when the Census has insufficient records about whether an address is valid and nonduplicated; the Census then imputes the status of the unit.

Assumptions

1. During the interval between the census taking and the follow-up ACE, the proportions of people that were surveyed in the census and those who weren't remain the same. This means no new people are born, die, immigrate, or emigrate.
2. The two surveys are independent, meaning the chance that a person was counted in the ACE does not change depending on whether or not they were counted in the census.
3. There is sufficient time between the census and the ACE for all "marked" individuals to be randomly dispersed in the population.

Design

After the initial census period, the Census Bureau should conduct a postcensus population estimate, currently called the Accuracy and Coverage Evaluation (ACE), using a mark and recapture method to account for the people missed in the original count. In this mark and recapture method, a second sample should be taken of randomly selected areas within each state. The extent of the overlap should then used to adjust the census population estimates for each state. The formula used for adjustment is:

---

\[ \frac{R}{M} = \frac{C}{N} \]

and therefore, \[ N = \frac{CM}{R} \]

where

- \( N \): Population size for each state
- \( M \): Total number of individuals surveyed and recorded in the census
- \( C \): Total number of individuals surveyed during the ACE
- \( R \): Number of individuals surveyed in the census that were then resurveyed during the ACE.\(^7\)

**Error Analysis**

While the mark and recapture method would improve upon the raw census data, there will still be some error. This can be found using the standard error:

\[ SE = \sqrt{\frac{(M+1)(C-1)(M-1)(C-1)}{(R-1)(R+2)}} \]

The standard error can be used to find a 95% confidence interval for the population size, where

\[ \text{Population} = N \pm (1.96)(SE) \]

This formula gives a range where 95% of the time, when calculated using sample data, the actual value of the population is included.\(^8\)

**Testing the Model**

In order to find and compare the error of this model with the undercounts of the census, we would need data from both the 2010 Census and ACE. So, since the error for the 2010 Census cannot yet be calculated, to test our model we compared the standard error of the mark and recapture method in the 2000 ACE to the 2000 Census. Using data from the Statistical Research Center of the Census Bureau, the standard error of the ACE was found to be 541,631. This error is less than the number of people undercounted in the 2000 Census, 1,331,656 people.\(^9\) This means that using the mark and recapture method will yield less error than using the census alone.

---

8. [http://www.neiu.edu/~jkasmer/Biol380/Labs/mark&.htm](http://www.neiu.edu/~jkasmer/Biol380/Labs/mark&.htm)
**APPORTIONMENT**

**Rationale**

To understand our choice for apportionment, we must first understand a bit about the basics of apportionment. The term "standard divisor" refers to the ratio of the total population to the number of seats, while the "standard quota" for a state refers to the population of the state divided by the standard divisor. This produces a number which will be used to calculate the number of seats for a state; a state's "upper" and "lower" quotas are the closest integers above and below this number.

In choosing a method, we faced a dilemma: choose a method that would allow for paradoxes when population, the number of seats, or the number of states changed, or choose a method that would violate the quota rule, where each state must receive either their upper or lower quotas. Unfortunately, both cannot be satisfied; the Balinski–Young impossibility theorem says that both the paradoxes and the violated quota rule cannot be avoided.

**Assumptions**

1. No new states will be added to the union; therefore, the *New States* paradox (adding a new state but keeping the number of representatives constant resulting in an increase in representatives for an existing state) would never occur.
2. The total number of seats in the House of Representatives will remain constant. Since the number of seats is not increasing, the *Alabama* paradox (a state losing a seat when the total number of seats increases) will never occur.
3. Each state must have at least one representative.

**Design**

For apportionment, the CAWES model is proposed to determine the number of representatives each state receives in the House of Representatives. Because we have limited the number of seats and number of states, the only paradox worth noting is the population paradox, where if one state has an increase in population at a faster rate than others, it may still lose seats. The CAWES model accounts for the population paradox and does not violate the quota rule, which states that the apportionment always allocates only lower or upper bounds of the quota.

The CAWES Model attempts to avoid as many of the paradoxes as possible. By its definition, it avoids breaking the quota rule. All states are given the lower quota, and then one more is added only to states that qualify, giving them their upper quotas—nothing will ever fall outside these bounds. As for paradoxes, our model avoids the population paradox by awarding extra seats to those that are increasing at the fastest rate. None of the other paradoxes are possible in this scenario either, because neither states nor seats are being added in our model. This is because none have been changed since the sixties, and none are planning on changing anytime soon. A bonus of our model is that it takes into account not only the present, but also the future. By predicting the rate of change of each state, we can better represent the state not only at the time of the census, but also between censuses.
Variables:

- $P_{\text{total}}$: Total population of the United States (in 2010)
- $P_{2005}$: the Census projected population for 2005
- $P_{2015}$: the Census projected population for 2015
- $M$: Number of seats in the House of Representatives
- $P_i$: Population of an individual state (where $1 \leq i \leq 50$)
- $SD$: Standard divisor
- $Q_i$: Standard quota for each state
- $LQ_i$: Lower quota for each state
- $Y$: Seats given to states with an $LQ_i$ of zero
- $S$: Surplus number of seats
- $R_i$: Number of representatives each state receives
- $r_i$: Rate of population increase for each state

Model

First, the total population of the United States is divided by the number of seats in the House of Representatives to calculate the standard divisor:

$$SD = \frac{P_{\text{total}}}{M}$$

Each state's population is divided by the standard divisor to calculate the standard quota, or the number of representatives each state is expected to receive:

$$Q_i = \frac{P_i}{SD}$$

The lower quota of each state is calculated by rounding the quota down to the nearest whole number. Those states with a lower quota of zero automatically get one seat added so that each state will have at least one representative. Next, the surplus is calculated by subtracting the sum of all the lower quotas and those given to states with a zero for a lower quota from the total number of representatives (435):

$$S = M - \sum_{i=1}^{50} LQ_i - Y$$

Then, the rate of population increase for each state in 2010 is also calculated from the census projected population rates from 2005 and 2015 such that

$$r_i = \frac{P_{2005} + P_{2015}}{10}$$

The top $S$ number of states with the highest rate of population increase in 2010 will also receive one additional seat unless that state's $LQ_i$ was originally zero. Thus those states that have a population increasing more than others will receive extra seats and therefore prevent the population paradox. If done correctly, the total number of representatives should equal the number of representatives to be divided:

$$\sum_{i=1}^{50} R_i = M = 435.$$
<table>
<thead>
<tr>
<th>State</th>
<th>2005 Pop.</th>
<th>2010 Pop.</th>
<th>2015 Pop.</th>
<th>Standard Quota</th>
<th>Rate of Change</th>
<th>Lower Quota</th>
<th>2010 Predicted Apportionment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>4447100</td>
<td>4527166</td>
<td>4596330</td>
<td>6.68</td>
<td>0.0999</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Alaska</td>
<td>626932</td>
<td>661110</td>
<td>694109</td>
<td>0.98</td>
<td>0.0999</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Arizona</td>
<td>5130632</td>
<td>5868004</td>
<td>6637381</td>
<td>8.65</td>
<td>0.1003</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Arkansas</td>
<td>2673400</td>
<td>2777007</td>
<td>2875039</td>
<td>4.10</td>
<td>0.0999</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>California</td>
<td>3387148</td>
<td>3603859</td>
<td>3806714</td>
<td>53.15</td>
<td>0.0998</td>
<td>53</td>
<td>54</td>
</tr>
<tr>
<td>Colorado</td>
<td>4301261</td>
<td>4617962</td>
<td>4831554</td>
<td>6.81</td>
<td>0.0989</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Connecticut</td>
<td>3405565</td>
<td>3503185</td>
<td>3577490</td>
<td>5.17</td>
<td>0.0997</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Delaware</td>
<td>783600</td>
<td>836687</td>
<td>884342</td>
<td>1.23</td>
<td>0.0997</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Florida</td>
<td>15982378</td>
<td>17509827</td>
<td>19251691</td>
<td>25.82</td>
<td>0.1006</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>Georgia</td>
<td>8186453</td>
<td>8925796</td>
<td>9589080</td>
<td>13.16</td>
<td>0.0996</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Hawai‘i</td>
<td>1211537</td>
<td>1276552</td>
<td>1340674</td>
<td>1.88</td>
<td>0.1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Idaho</td>
<td>1293953</td>
<td>1407060</td>
<td>1517291</td>
<td>2.08</td>
<td>0.0999</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Illinois</td>
<td>12419293</td>
<td>12699336</td>
<td>12916894</td>
<td>18.73</td>
<td>0.0998</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Indiana</td>
<td>6080485</td>
<td>6249617</td>
<td>6392139</td>
<td>9.22</td>
<td>0.0998</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Iowa</td>
<td>2926324</td>
<td>2973700</td>
<td>3009907</td>
<td>4.39</td>
<td>0.0998</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Kansas</td>
<td>2688418</td>
<td>2751509</td>
<td>2805470</td>
<td>4.06</td>
<td>0.0998</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Kentucky</td>
<td>4041769</td>
<td>4163360</td>
<td>4265117</td>
<td>6.14</td>
<td>0.0998</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Louisiana</td>
<td>4468976</td>
<td>4534310</td>
<td>4612679</td>
<td>6.69</td>
<td>0.1001</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Maine</td>
<td>1274923</td>
<td>1318557</td>
<td>1357134</td>
<td>1.94</td>
<td>0.0998</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Maryland</td>
<td>5296486</td>
<td>5600563</td>
<td>5904970</td>
<td>8.26</td>
<td>0.1</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>6349097</td>
<td>6518868</td>
<td>6649441</td>
<td>9.61</td>
<td>0.0997</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Michigan</td>
<td>9938444</td>
<td>10207421</td>
<td>10428683</td>
<td>15.05</td>
<td>0.0998</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Minnesota</td>
<td>4919479</td>
<td>5174743</td>
<td>5420636</td>
<td>7.63</td>
<td>0.0999</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Mississippi</td>
<td>2844658</td>
<td>2915696</td>
<td>2971412</td>
<td>4.30</td>
<td>0.0997</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Missouri</td>
<td>5595211</td>
<td>5765166</td>
<td>5922078</td>
<td>8.50</td>
<td>0.0999</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Montana</td>
<td>902195</td>
<td>933005</td>
<td>968598</td>
<td>1.38</td>
<td>0.1003</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nebraska</td>
<td>1711263</td>
<td>1744370</td>
<td>1768997</td>
<td>2.57</td>
<td>0.0998</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Nevada</td>
<td>1998257</td>
<td>2352086</td>
<td>2690531</td>
<td>3.47</td>
<td>0.0997</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>1235786</td>
<td>1314821</td>
<td>1385560</td>
<td>1.94</td>
<td>0.0997</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>New Jersey</td>
<td>8414350</td>
<td>8745279</td>
<td>9018231</td>
<td>12.90</td>
<td>0.0997</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>New Mexico</td>
<td>1819046</td>
<td>1902057</td>
<td>1980225</td>
<td>2.81</td>
<td>0.0999</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>New York</td>
<td>18976457</td>
<td>19258082</td>
<td>19443672</td>
<td>28.40</td>
<td>0.0998</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>North Carolina</td>
<td>8049313</td>
<td>8702410</td>
<td>9345823</td>
<td>12.83</td>
<td>0.0999</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>North Dakota</td>
<td>642200</td>
<td>635468</td>
<td>636623</td>
<td>0.94</td>
<td>0.1006</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Ohio</td>
<td>11353140</td>
<td>11477557</td>
<td>11576181</td>
<td>16.93</td>
<td>0.0999</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>3450654</td>
<td>3521379</td>
<td>3591516</td>
<td>5.19</td>
<td>0.1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Oregon</td>
<td>3421399</td>
<td>3596083</td>
<td>3790996</td>
<td>5.30</td>
<td>0.1003</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>12281054</td>
<td>12426603</td>
<td>12584487</td>
<td>18.33</td>
<td>0.1001</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>1048319</td>
<td>1086575</td>
<td>1116652</td>
<td>1.60</td>
<td>0.0996</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
South Carolina 4012012 4239310 4446704 6.25 0.0998 6 7
South Dakota 754844 771803 786399 1.14 0.0999 1 1
Tennessee 5689283 5965317 6230852 8.80 0.0999 8 8
Texas 20851820 22775044 24648888 33.59 0.0999 33 34
Utah 2233169 2417998 2595013 3.57 0.0998 3 4
Vermont 608827 630979 652512 0.93 0.1 0 1
Virginia 7078515 7552581 8010245 11.14 0.0999 11 11
Washington 5894121 6204632 6541963 9.15 0.1002 9 10
West Virginia 1808344 1818887 1829141 2.68 0.1 2 2
Wisconsin 5363675 5554343 5727426 8.19 0.0998 8 8
Wyoming 493782 507268 519886 0.75 0.0999 0 1
TOTAL 280849847 294955998 308405796 435

**Figure 1:** The CAWES Model with the predicted values for 2005, 2010, and 2015.

States that gain more than one seat (using the CAWES Model as opposed to the currently used Huntington–Hill model):
- Texas

States that gain one seat:
- Arizona
- Florida
- Hawai‘i
- Maine
- Nebraska
- New Mexico
- North Carolina
- Tennessee
- West Virginia

States that lose one seat:
- Arkansas
- California
- Idaho
- Maryland
- Michigan
- Minnesota
- Missouri
- North Dakota
- Ohio
- Oregon
- Washington

As we can see, most of the data fits the current trend that most of the population is shifting from the North and East to the South and West. Note that just because a state loses a seat from the current model to the CAWES model does not necessarily mean that that state is losing citizens. Rather, they are either growing at a slower rate than the rest of the union or they have been over-represented in the past.
CONGRESSIONAL DISTRICTS

Rationale

When the demographics of an area change, redistricting occurs so that the population is equally distributed among the districts and each representative represents roughly the same percentage of the state. Currently for thirty-six states, the state legislature has the primary responsibility for creating a redistricting plan for its given state. In some states this is subject to approval by the state governor. Five states redistrict under the direction of a commission. In two states, state legislatures must approve redistricting plans proposed by independent bodies. Seven states have only a single representative for the entire state.

Gerrymandering refers to redistricting in which electoral districts are deliberately modified to give advantage to a particular group or candidate; this is usually done by linking similar areas, that contain a particular group of people, in a contorted or unusual shape. This can be prevented through legislation. Currently certain types of gerrymandering are illegal, such as those that deliberately isolate a certain race. Gerrymandering is a problem because it does not reflect the actual will of the entire population through true democracy.

Assumptions

1. There are two opponents in redistricting, the Republicans and the Democrats.
2. The districts should be fairly drawn and should prevent gerrymandering.

Design

The redistricting process could be executed fairly if each state were to use the our modified version of the Zeph–Landau method for distributing population size, which is based on the philosophy of "I cut, you choose" used when dividing a cake. Thus, we recommend the modified Zeph–Landau method for reapportionment.

This method works by dividing the state into 25 sections such that all the sections have an equal number of people. For example, see Figure 2.
Next, a random line is drawn dividing the state in two (see Figure 3). The placement of this line does not matter.

Step 1: The Republican Party can divide piece #1 of the split into \( k \) districts and have the Democratic Party divide piece #2 into \( n - k \) districts (where \( n \) is the total number of districts allocated by the reapportionment). The resulting wins and losses for each party, based on the known voting records for these districts, are recorded for this division.

Step 2: Then, the Democratic Party divides piece #1 of the split into \( k \) districts and have the Republican Party divide piece #2 into \( n - k \) districts. These resulting wins and losses for each party, based on the known voting records for these districts, are recorded for this division.

---

Figure 4: The numbers in each section represent the ratio of Republicans to Democrats based on polls. The first diagram shown is Step 1, in which the Republican Party divides the left piece and the Democratic Party divides the right piece. The second diagram is Step 2, in which the Democratic Party divides the left piece and the Republican Party divides the right piece. 

After repeating these procedures multiple times with different dividing lines, a table is constructed with this gathered data as shown below. The table is ordered by ascending wins for one of the Steps. There will be a point at which the favorable step for that party switches between Step 1 and Step 2. This is the point in which the number of representative districts is about equal toward each party.

<table>
<thead>
<tr>
<th>Line</th>
<th>Step 1: Part 1+Part 2 = total</th>
<th>Step 2: Part 1+Part 2 = total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line 1</td>
<td>1+1=2</td>
<td>1+3=4</td>
</tr>
<tr>
<td>Line 2</td>
<td>1+1=2</td>
<td>1+2=3</td>
</tr>
<tr>
<td>Line 3</td>
<td>3+0=3</td>
<td>1+1=2</td>
</tr>
<tr>
<td>Line 4</td>
<td>4+0=4</td>
<td>1+0=1</td>
</tr>
</tbody>
</table>

Figure 5: This figure shows an example of several divisions (referred to as Line 1, Line 2, etc.) and the resulting Republican Party win distribution. For example, Figure 3 would be the Line 3 division which in Step 1 gives Republicans 3 wins while Step 2 gives Republicans 2 wins. In this example, the Republican Party would change from Step 2 to Step 1 between Line 2 and Line 3.

The change between the steps serves as a point in which the party distribution is about equal. This leaves four possible rearrangements of the distribution. For example, in Figure 5, Republicans can have:

The modified Zeph–Landau method would then look to the party distribution for Democrats and Republicans in that state, and the party with the most supporters would get the most districts in that state. For example, in the above scenario, the given state is majority Republican and thus the Republicans would end with 3 districts. This would narrow the choices down to:

- Step 1 and Line 3, leaving Republicans with 3 districts.
- Step 2 and Line 2, leaving Republicans with 3 districts.

Between these two choices, the result would be randomly chosen.

The reason the modified Zeph–Landau method is used rather than the Zeph–Landau method is because the latter method does not account for states that lean heavily toward a party but assumes that a gerrymandering-free state is one that is equally Republican and equally Democrat. The modified Zeph–Landau method, however, allows states that are not equally distributed to have slanted districts, yet not so slanted that the districts consistently dominate one another.

**Testing.** This model could be applied in every state. Because there is no current system to prevent gerrymandering other than legislative agreement, the modified Zeph–Landau method would be ideal.

**CONCLUSIONS**

Our population estimates error using the mark and recapture method was estimated at approximately half a million people, far less than the 1.3 million people uncounted by the Census. While our method is not perfect, it is far more accurate than using the Census alone.

The Huntington–Hill method is currently used to apportion the seats in the House of Representatives to the states. This method has a major problem—it can violate the quota method, causing a state to get far more or fewer states than would logically be expected. Our CAWES model does not have this problem; we assign all states at least the minimum number of seats they would expect to receive and then assign the remaining seats based on the rate the population of each state is increasing. Since the states with the fastest increasing population are more likely to "deserve" seats before the next apportionment, we decided to assign the extra seats to these states.

Compared to the current Huntington–Hill method of apportionment, our CAWES model takes into account not just the current population, but also the predicted future population. This provides for a better representation of the state into the future rather than just at the present. It also doesn't severely damage the amount of seats according to the current model that predicts the 2010 data.

We used a modified version of the Zeph–Landau method to redistribute states. This stems from the philosophy "I cut, you choose" and makes both parties work together in order to divide the districts evenly among the parties. In this way, one district can never consistently dominate the others.
The models proposed are the best that we recommend but have flaws nevertheless. The mark and recapture method is a statistical analysis and therefore will always contain some degree of error. For example, if the time between surveying is too great, then the demographics may have changed, and therefore the two surveys would not be comparable. The CAWES model becomes potentially problematic upon introduction of the Alabama and New States paradoxes, but since new states or new seats cannot be added, this is not a problem. Also, since the seats are allocated based on the rate of population change, the Population Paradox should also not be a problem. The modified Zeph–Landau method does not work for more than two parties and takes longer than the current methods of redistricting because it requires every district to be reevaluated after every Census.

Another idea for a possible future model to impute the census is to analyze past birth, death, immigration, and emigration data on a state-by-state basis to calculate the states' net population changes. As long as good records are kept, this would allow for a census to be approximated any year and therefore assist in census imputation. This model would not account for illegal immigrants.
REFERENCES


Recent court decisions about the census, adjusting for census undercount and the use of census data to apportion Congress and the Electoral College [Word document]. (n.d.). Retrieved from University of Michigan website: http://www.psc.isr.umich.edu/dis/acs/handouts/SupremeCourt.doc