

Impartial and Unbiased Apportionment: Meeting the Ideal of One Person, One Vote

Ross Hyman*

Cardinal Intellectual Property, 1603 Orrington Avenue, Evanston IL 60201

Nicolaus Tideman[†]

Economics Department, Virginia Tech, Blacksburg VA 24060

(Dated: January 31, 2022)

Abstract

We develop two classes of Congressional apportionment methods that we call impartial methods and unbiased methods. Both classes of methods are motivated by grouping states into “families,” sets of states with divisor-method quotas that round down to the same integer. Impartial methods apportion the same number of seats to families of states containing the same total population, whether a family consists of a large number of small-population states or a small number of large-population states. Impartial apportionment methods are divisor methods applied to families of states instead of individual states. We show that the Webster method applied to families, but not the Huntington Hill method applied to families, is immune to the Alabama paradox and that no apportionment method can be both impartial and immune to the New State paradox. Unbiased methods apportion seats so that if states are drawn repeatedly from the same distribution, then the expected value of the total number of seats apportioned to each family equals the expected value of the total divisor-method quota for that family. Unbiased apportionment methods avoid the Alabama paradox and the New State paradox, yet they are not divisor methods because they are not homogeneous. There is a different unbiased apportionment method associated with each distribution of state populations. We develop a model of population growth and migration leading to a log normal distribution of state populations, and we determine the unbiased apportionment method for the log-normal-distribution.

I. INTRODUCTION

Every ten years, the U.S. House of Representatives is reapportioned according to the census for that decade. The Constitution specifies that “Representatives shall be apportioned among the several States according to their respective numbers” but prescribes no method to accomplish this. It is not a simple matter of multiplying the total number of seats in the House by each state’s fraction of the population, since the resulting numbers of seats will generally not be integers. Many apportionment methods that assign each state an integer number of seats have been proposed, including methods proposed by Thomas Jefferson, John Quincy Adams, Alexander Hamilton, and Daniel Webster.^{1,6,7} Due to the political implications of the gain or loss of even one seat, disputes over apportionment methods have occurred many times, on the grounds that a proposed method is biased against states with small or large populations. The first such dispute resulted in George Washington’s first veto.¹ The last such dispute, over the apportionment for the 1990 census, went to the Supreme Court.³

We present an alternative way to think about apportionment, in which the fundamental entities are not individual states but “families” of states and we introduce impartial apportionment methods and unbiased apportionment methods. Impartial methods apportion seats among families of states according to their respective populations, regardless of whether a family contains a large number of small-population states or a small number of large-population states. Unbiased methods apportion seats so that if states are drawn repeatedly from the same distribution, then the expected value of the total number of seats apportioned to each family equals the expected value of the total divisor-method quota for that family.

In sections II and III we present a history of methods used or proposed for United States congressional apportionment, describe the Alabama paradox and the New State paradox, motivate apportionment methods using a “slide rule,” and show that divisor methods applied to states are not susceptible to the Alabama paradox or the New State paradox. In sections IV - VII we define families of states and show that the Webster and Huntington-Hill methods applied to families are impartial and susceptible to the New State paradox. We find that Webster’s method applied to families of states avoids the Alabama paradox while Huntington-Hill applied to families is susceptible to it. In sections VIII - XII we consider state populations that are drawn from a specified probability distribution. We determine,

for any probability distribution for populations, the apportionment method associated with that distribution that, when applied to states, is unbiased and avoids the Alabama and New State paradoxes. We develop a model of population growth and migration leading to a log normal distribution of state populations and determine the unbiased apportionment methods for the states drawn from the log-normal distribution.

II. A BRIEF HISTORY OF APPORTIONMENT

Article I, section 2 of the U.S. Constitution specifies:

Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective Numbers, . . . The actual 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. The Number of Representatives shall not exceed one for every thirty Thousand, . . .

The Constitution does not specify a method of rounding the fractions that occur when apportioning integer numbers of representatives among states whose populations are not integer multiples of any feasible district size.

When Congress received the results of the 1790 census, it debated several methods of apportionment before settling on “Hamilton’s method,” which involves deciding a size of the House, allocating seats including fractions to all states, rounding down to integers, and then apportioning the resulting extra seats to the states with the largest fractional remainders.^{1,2} George Washington vetoed the bill that Congress passed, the only veto of his first term, on the ground that it gave Connecticut one representative for every 29,605 persons counted in the Census, in violation of the Constitutional requirement that the representation of each state not exceed one for each 30,000.^{1,2} Congress responded by adopting “Jefferson’s method,” which involved dividing each state’s population by 33,000 and rounding down to an integer.^{1,2}

Jefferson’s method was used for apportionments for all censuses through 1830, but Congress grew concerned over the method’s bias in favor of large-population states. In the debate surrounding the apportionment for the 1830 census, other possibilities began to be

discussed. John Quincy Adams proposed rounding up rather than down. James Dean, a mathematics professor from Vermont, proposed that the decision about whether a given state's allotment should be rounded up or down should depend on which produced a ratio of persons per representative closer to the overall target, and Daniel Webster proposed ordinary rounding. For the 1840 census, Congress chose Webster's method.^{1,2}

For the apportionments from 1850 to 1900, Congress described the apportionment it chose as being Hamilton's method, which had been their first choice in 1790, before Washington's veto. However, in 1850, 1870, 1880, and 1890, Congress deliberately chose a size for the House such that Hamilton's method and Webster's method yielded the same apportionment.^{1,2}

For the apportionments of 1910 and 1930, Congress went back to Webster's method. (Congress failed to pass a new apportionment in response to the census of 1920.) For the apportionment in response to the 1930 census, Congress chose to keep the House size at 435, knowing that at that size, Webster's method agreed with the Huntington-Hill method for the 1930 census. The Huntington-Hill method rounds down or up depending on whether a state's ideally assigned number of seats, a real number that is generally not an integer, is less than or greater than $\sqrt{S(S+1)}$, where S is the integer obtained by rounding the assigned number of seats down. Since 1940, Congress has chosen to use the Huntington-Hill apportionment method, keeping the size of the house at 435 seats.^{1,2}

In 1880, when the Census Bureau calculated the allotments under Hamilton's method for every House size from 275 to 350, Congress became aware that Hamilton's method is subject to the "Alabama paradox." This is the phenomenon that when the size of the House increases, it is possible for a state to lose a seat.^{1,2}

In 1907 Oklahoma became a state, and the 'New State paradox' was discovered when the Census Bureau calculated allotments using Hamilton's method. The New State paradox occurs when the size of the House increases through the addition of a new state along with additional seats appropriate for the size of the new state, and an existing state loses a seat.^{1,2}

Michel Balinski and H. Peyton Young¹ identified apportionment methods that avoid both the Alabama paradox and the New State paradox. They also showed that if the populations of states are drawn from a uniform distribution, then a) the Huntington-Hill method,^{4,5} which Congress has used since 1930, is biased in favor of states with small populations and b) Webster's method does not favor states with small or large populations. In addition, they

applied bias criteria to apportionments based on the Webster and Huntington-Hill methods and noted that the Webster apportionments showed less bias.¹

For the 1990 census, the state of Massachusetts, which would have gained a seat with Webster’s method, sued the U.S government on the grounds that the bias in favor of small-population states identified by Balinski and Young made the use of the Huntington-Hill method for congressional apportionment unconstitutional. The briefs in favor of the government’s position were written by Laurence Ernst, a Census Bureau mathematician, who showed that for every criterion and measure of bias for which Webster apportionments were the least biased, there was a different criterion, for which the Huntington-Hill apportionments were least biased. The court decided in favor of the federal government, holding that, as there was no objective way to choose between the contending bias criteria, the choice between these contending methods was entirely subjective and within the purview of the Congress. Therefore, continued use of the Huntington-Hill method for congressional apportionment was constitutional.³

The problem of apportioning seats in a legislature to parties under a system of proportional representation is mathematically the same as the problem of apportioning seats in Congress among states. For a discussion of issues of apportionment in the context of proportional representation, see Pukelsheim.⁹

III. THE APPORTIONMENT SLIDE RULE

The integer number of seats apportioned to state c , s_c , would be perfectly proportional to the population of the state, v_c , if there were an intended district size D such that $s_c = v_c/D$ for every state. If the populations of states came only in multiples of the intended district size D , then a state would be assigned one seat for each D persons, and there would be no apportionment problem. The apportionment problem arises because the “quota” for state c , $q_c = v_c/D$, is generally not an integer. We define the quota as the divisor-method quota in terms of the intended district size D , which can be any positive real number, and not necessarily the average district size v_T/s_T where v_T is the total population and s_T is the total number of seats. The quota using D and not the quota using v_T/s_T must be used for the Alabama paradox to be avoided.

Apportionment can be understood in the following visual way. Imagine two sliding rulers,

one on top of the other, with logarithmic scaling on each. Let the top ruler be the Population Ruler and the bottom ruler be the Seats Ruler. Logarithmic scaling ensures that sliding the rulers corresponds to changing the common ratio, D , of population to unrounded seats. For a given value of D , there is a relative positioning of the two rulers, such that the quota of seats allotted to a state with a given population is the number of seats on the Seats Ruler directly below the population of the state on the Population Ruler. In general, the specified number of seats will not be an integer. Most apportionment methods partition the Seats Ruler into integer seat segments. A rounding mark, $r(f)$, is placed on the Seats Ruler somewhere between integers f and $f + 1$. The segment of the Seats Ruler between consecutive rounding marks $r(f - 1)$ and $r(f)$ is the integer seat region for f seats. When the Population Ruler is positioned over the Seats Ruler so that v_c is over any part of the f seat region, that is $r(f - 1) \leq q_c < r(f)$, state c is assigned f seats.

In general, rounding marks can be dependent on f , D , and the state populations $\{v\}$, and can move if any of these change. These movements are the cause of the various paradoxes of apportionment. For example, if D is decreased (increased) the Population Ruler moves to the right (left). If a rounding mark dependent on D moves to the right (left) faster than the Population Ruler so that it outruns a state, then that state loses (gains) a seat, which is contrary to the motion of D , decreasing (increasing) the intended district size. This is the Alabama paradox. The total number of seats is adjusted by adjusting D , which corresponds to moving the Population Ruler relative to the Seats Ruler. For apportionment methods immune to the Alabama paradox, the total number of seats will not decrease (increase) when D decreases (increases).

If a state is added or removed, while D and the populations $\{v\}$ of the other states are held fixed, and a rounding mark dependent on state populations consequently moves through the population of an existing state, then that state's allotment will change. This is the New State paradox.

The Alabama paradox will never occur if $\partial \log r / \partial \log D \geq -1$ and the New State paradox will never occur if rounding for state c can depend on v_c , f , and D but not on the population of any other state. Both of these conditions are satisfied if $r(f)$ depends only on f and not on D or $\{v\}$. In other words, the rounding mark that determines the number of seats allotted to state c depends only on the integer part of q_c and not the quota of any other state. Such methods are called divisor methods¹. A result of our paper is that there are methods

that are immune to the Alabama paradox and the New State paradox yet are not divisor methods.

IV. FAMILIES OF STATES

One way to determine the population-size bias of an apportionment is to divide the states into groups of small- and large-population states and compare the average representation per person of the groups. However, care must be taken to draw the dividing line between small-population and large-population states at an integer on the Seats Ruler. To see this, consider the region of the Seats Ruler extending from any integer f to $f + 1$. For all apportionment methods for which a state cannot be allotted fewer seats than another state with a smaller quota, which includes all of the methods discussed in this paper, the states that are rounded down to f seats will have lower quotas than the states that are rounded up to $f + 1$ seats. If one placed the dividing line between the small and large states between two consecutive integers, rather than exactly on an integer, then the lower quota states of the divided range, which are rounded down, are grouped with the small states and the higher quota states of the divided range, which are rounded up, are grouped with the large states. Then if one compares the average number of representatives per person in the two groups, the apportionment will appear to be biased in favor of the large states. Therefore, if one is concerned about size bias over an entire apportionment, it is essential to keep states between consecutive integers on the Seats Ruler together.

In this section we reexamine apportionment by treating groups of states between consecutive integers on the Seats Ruler, rather than the individual states, as the fundamental entities that are to be apportioned seats.

All the states whose quotas $q_c = v_c/D$ fall in the region $f \leq q_c < f + 1$, that is, all states with quotas whose integer part is f , we call the f family of states. We call the sum of the quotas of all states in the f family Q_f . In general, Q_f is not an integer. We round Q_f to an integer following a prescribed rounding method to determine the number of seats for the f family, and we call that integer S_f . The value of S_f depends on the apportionment method.

Once an integer S_f is chosen, assigning seats to the N_f states in the f family is uniquely

determined by

$$\begin{aligned} M_f &= (f + 1)N_f - S_f \\ M_{f+1} &= S_f - fN_f \end{aligned} \tag{1}$$

where the M_f smallest states in the family are assigned f seats each and the M_{f+1} largest states in the family are assigned $f+1$ seats each. Which states are in which group is uniquely determined by insisting that $s_c \geq s_d$ if $v_c > v_d$.

There is no other grouping of states, besides the family groupings, that can contain a variable number of states and for which the numbers of seats assigned to individual states in the group are uniquely determined by the number of seats assigned to the group as a whole.

The rounding mark in the f family, for family apportionment methods, is neither explicitly determined nor needed, but rather is any number from f to $f + 1$ that is smaller than the smallest state quota that is rounded up and larger than the largest state quota that is rounded down.

V. IMPARTIAL APPORTIONMENT

Apportionment methods for families can be characterized by the rounding rule used to convert real Q_f to integer S_f . In general, the rounding rule can depend on Q_f , f , D , and $\{v\}$. If we want the rounding rule to apportion the same numbers of seats to a family containing a large number of small-population states as it does to another family with the same total population, but containing a small number of large-population states, then the rounding rule can depend only on Q . Furthermore, we want the number of apportioned seats for the family to be as close as possible to Q according to a measure of closeness such as Webster rounding which minimizes $|S - Q|$ or Huntington-Hill rounding which minimizes $|\log S - \log Q|$. We call an apportionment method that satisfied these properties, such as Webster's method applied to families and the Huntington-Hill method applied to families, an impartial method. Other divisor methods as well as Hamilton's method, can also be made impartial by applying them to families rather than states.

Rounding family quotas will produce the same apportionment as rounding state quotas when there is no more than one state in a family. However, when there is more than one state in a family, rounding *family* quotas using any particular rounding rule may produce

TABLE I. Differences between Webster for families and Webster for states

State	Quota	Seats	
		Webster for Families	Webster for States
North Dakota	1.024	1	1
South Dakota	1.166	1	1
Delaware	1.302	1	1
Montana	1.426	1	1
Rhode Island	1.443	2	1
Maine	1.791	2	2
New Hampshire	1.812	2	2
Hawaii	1.918	2	2
1 Family	11.883	12	11
Louisiana	6.124	6	6
Alabama	6.608	6	7
South Carolina	6.733	7	7
6 Family	19.465	19	20
Minnesota	7.501	7	8
Colorado	7.596	8	8
Wisconsin	7.748	8	8
7 Family	22.846	23	24
Tennessee	9.087	9	9
Massachusetts	9.240	9	9
Arizona	9.405	10	9
9 Family	27.733	28	27

a different apportionment than rounding *state* quotas using that same rule. This occurs because summing and then rounding will in general produce a different result than rounding and then summing.

Table I shows Webster for families (summing and then rounding) and Webster for states (rounding and then summing) apportionments for 435 seat apportioned using 2020 census data for those families where the two apportionments differ.

To make the total number of seats in each family as close as possible to its family quota, Webster for Families apportions an additional seat to Rhode Island and Arizona and one fewer seat to Alabama and Minnesota than Webster for States. For each of the families involved, one can see that the difference between the Webster for families and Webster for states is due to a skewed distribution of states within those families. In the 1-Family, five states have quotas less than 1.5 while three states have quotas above 1.5. In the 6-Family, one state has a quota less than 6.5 and two states have quotas above 6.5. In the 7-Family, all three of its states have quotas above 7.5. In the 9-Family, all three of its states have quotas less than 9.5.

VI. THE ALABAMA PARADOX AND IMPARTIAL APPORTIONMENT

Divisor methods applied to states are immune to the Alabama paradox. But not all divisor methods applied to families are immune. This is because unlike states, families can change their population size as D is changed. This occurs when a state that is in one family for one value of D is in a different family for another value of D . There is no counterpart to D dependent population change for divisor methods applied to states, since the population of a state remains the same as D is changed.

When D is decreased, and therefore the Population Ruler is slowly moved to the right, a state leaves the f family and enters the $f + 1$ family just as its state quota is the integer $f + 1$ at the boundary between the f and $f + 1$ families. The family quota, Q_f , for the f family decreases by $f + 1$, an integer amount, and the family quota, Q_{f+1} , for the $f + 1$ family increases by $f + 1$. This can have effects on the other states in the two families if the rounded family quota, S , for a family changes by a different amount than the unrounded family quota, Q , when a state enters or leaves the family. This can happen if the rounding up or down of Q depends on the integer part of Q , as is the case for Huntington-Hill rounding. It cannot happen if the rounding up or down of Q does not depend on the integer part of Q , as is the case for Webster rounding.

Consider two states, one with quota 0.999 and the other with quota 1.43. Each is the only state in its family, the 0 and 1 families respectively. Using Huntington-Hill $\sqrt{S(S+1)}$ rounding on the family quota Q , the first state's 0.999 family quota is rounded up to 1 and the second state's 1.43 quota is rounded up to 2. As D is decreased so that every state's

quota gets magnified by $1001/999$, the first state's quota is now 1.001 and the second's quota is now 1.433. Now the two states are both in the 1 family with quota 2.433 which is rounded to 2. Therefore, the first state gets one seat as before, but now the second state also gets one seat. The second state lost a seat even though D decreased. This is the Alabama paradox.

If the rounding of Q depends on the integer part of Q , as with Huntington-Hill rounding, then the total number of states can go down, up, or stay the same as D decreases. In the above example, if the two states are the only states in the country, then the total number of seats went down from 3 to 2 as D decreased. For an example in which the total number of seats remains the same add a third state with quota initially 62.4375 which rounds down to 62 and the total number of seats is 65. After magnification by a factor of $1001/999$, the third state's quota is 62.5625 which rounds up to 63 and the total number of seats is also 65. In other words, if Huntington-Hill rounding is applied to family quotas to allocate 65 seats among three states with relative populations of 0.999, 1.43, and 62.4375, the allocation can be achieved either with the original quotas, yielding an apportionment of 1, 2, and 62, or by quotas scaled by $1001/999$, yielding an apportionment of 1, 1, and 63. This “multiple solution paradox” has not been mentioned previously in the literature because it cannot happen with apportionments by state.

The Alabama paradox, including the multiple solution paradox, cannot happen for Webster's method applied to families, for which the rounding up or down of the family quota does not depend on the integer part of the family quota. For conventional Webster rounding at 0.5, the first and second state are allotted one seat each, before and after D is changed, and no Alabama paradox occurs.

If combining states into families with the same integer part of their state quotas is worth considering, why not also combine families together that have the same integer part of their family quota, and so on? One disadvantage of combining families is that such methods can violate the Alabama paradox, and therefore can have more than one apportionment for the same total number of seats. This happens because rounding then summing is different than summing then rounding. To see this, consider three state quotas: 0.99999, 1.7, and 2.6. They are each in a different family. And they are rounded to 1, 2, and 3. Decrease D so the 0.99999 state quota gets moved to 1.0 and is now in the 1 family. The quotas of the 1 family and the 2 family are now approximately 2.7 and 2.6 respectively. If we stop here, the family quotas round to 3 and 3, the states are assigned 1, 2, and 3 seats respectively,

and no Alabama paradox has occurred. If we combine the 1 and 2 families into a family of the families with family quotas that round down to 2, then, this family-of-families has a quota of $2.6 + 2.7 = 5.3$, which rounds down to 5. The two families which compose the family-of-families are assigned 3 seats and 2 seats respectively, and the states are assigned 1, 2, and 2 seats. The Alabama paradox has occurred because a state has lost a seat when D has decreased.

VII. THE NEW STATE PARADOX AND IMPARTIAL APPORTIONMENT

Webster applied to families produces the least biased apportionment for a particular set of states, where the least biased apportionment is defined as the apportionment that minimizes $|S_f - Q_f|$ for every family. The method's rounding marks are determined impartially, dependent only on Q and therefore $\{v\}$. On the other hand, rounding marks of apportionment methods that are immune to the New State paradox cannot depend on $\{v\}$.

For example, a state with quota 2.6 that is the only state in the 2 Family is apportioned 3 seats and a state with quota 5.3 that is the only state in the 5 family is apportioned 5 seats. Add a state with quota 2.7 without changing D . The new family quota of the 2 family is 5.3 which rounds down to 5. Webster for families impartially apportions 5 seats to the 2 family and 5 seats to the 5 family. The 2.6 state is apportioned 2 seats and the 2.7 state is apportioned 3 seats. The 2.6 state has lost a seat to maintain impartiality between one large population state with a quota of 5.3 and two small population states with total quota 5.3. If one instead apportioned in a manner immune to the New State paradox, a state with quota 5.3 is rounded down to 5 and states with quotas 2.6 and 2.7 are rounded up to 3 seats each, which is biased in favor of the small population states. An apportionment method cannot be both impartial and immune to the New State paradox.

VIII. UNBIASED APPORTIONMENT

If the states are drawn from a known or theoretical distribution of states, then it is possible to base the apportionment method on the distribution of states rather than a particular sample. Using a distribution from which states are drawn, our task is to find the apportionment method that is unbiased in the sense that if states are repeatedly drawn from

the same distribution, then seats are apportioned such that $\langle S_f \rangle$, the expected value of the number of seats apportioned to the f family, is equal to $\langle Q_f \rangle$, the expected value of the total f family quota. Such an apportionment method, for which $\langle S_f \rangle = \langle Q_f \rangle$ for every f , we call an unbiased apportionment.

Balinski and Young¹ developed implications of state populations drawn from a uniform distribution. We generalize Balinski and Young's work by developing implications of a variety of population distributions. We do not adopt Balinski and Young's homogeneity assumption. The homogeneity assumption is that multiplying all state populations by a scale factor will not change an apportionment for the same total number of seats. To demonstrate that this apparently reasonable assumption is too restrictive, consider an apportionment method designed to be unbiased for states drawn from a bell-shaped distribution of population sizes. A set of states that are all drawn from the left end of the distribution should have rounding marks to the right of the midpoints of integer intervals to achieve unbiased apportionments, while a set of states with the same relative populations that are drawn from the right end of the distribution should have rounding marks to the left of the midpoints of integer intervals to achieve unbiased apportionments, violating homogeneity. It is because of this dropping of the homogeneity assumption that unbiased apportionment methods can be immune to both the New State paradox and Alabama paradox and yet, in general, not be divisor methods.

Define $p(v)$ as the density function for state population, v . We treat $p(v)$ as defined for all positive real numbers, even though population must in fact be an integer. Given a target district population, D , the average quota for the f family is

$$\langle Q_f \rangle = \int_{fD}^{(f+1)D} p(v) \frac{v}{D} dv, \quad (2)$$

and the expected value of the number of seats for the f family is

$$\langle S_f \rangle = f \int_{fD}^{rD} p(v) dv + (f+1) \int_{rD}^{(f+1)D} p(v) dv. \quad (3)$$

An apportionment is unbiased within the f family if

$$\langle S_f \rangle - \langle Q_f \rangle = f \int_{fD}^{rD} p(v) dv + (f+1) \int_{rD}^{(f+1)D} p(v) dv - \int_{fD}^{(f+1)D} p(v) \frac{v}{D} dv = 0. \quad (4)$$

In terms of the cumulative distribution function, $I(v)$, defined up to a constant by

$$\frac{dI(v)}{dv} = p(v), \quad (5)$$

we have

$$\begin{aligned}
\langle S_F \rangle - \langle Q_F \rangle &= (f+1)I((f+1)D) - fI(fD) - I(rD) - \int_{fD}^{(f+1)D} p(v) \frac{v}{D} dv \\
&= \int_{fD}^{(f+1)D} \frac{d}{dv} \left(I(v) \frac{v}{D} \right) dv - I(rD) - \int_{fD}^{(f+1)D} P(v) \frac{v}{D} dv \\
&= \frac{1}{D} \int_{fD}^{(f+1)D} I(v) dv - I(rD).
\end{aligned} \tag{6}$$

The condition for an unbiased apportionment is

$$I(rD) = \frac{1}{D} \int_{fD}^{(f+1)D} I(v) dv \tag{7}$$

so that

$$r(f, D) = \frac{1}{D} I^{-1} \left(\frac{1}{D} \int_{fD}^{(f+1)D} I(v) dv \right). \tag{8}$$

The dependence of the rounding marks on D mean that unbiased methods are not, in general, homogeneous divisor methods.

Since the rounding marks for unbiased apportionment are determined by an assumed state population distribution from which the states are drawn and not the actual state populations, it will, in general, differ from impartial apportionment, which is determined by the actual state populations and not by an assumed state population distribution. An apportionment from an unbiased method will not in general minimize $|S_F - Q_F|$ for each sample drawn from the distribution, and therefore will not be an impartial apportionment. Likewise, Webster for families applied to sets of states repeatedly drawn from the same distribution will not in general satisfy $\langle S_F \rangle = \langle Q_F \rangle$. However, impartial and unbiased apportionments will tend to agree if the number of states is large enough that any particular sample of states closely resembles the distribution it is drawn from.

IX. UNBIASED APPORTIONMENTS AND PARADOXES

The New State paradox will not occur if the rounding marks, which depend on the distribution from which the states are drawn, do not change when states are added to or removed from the sample. However, if properties of the distribution, such as its mean and standard deviation, are determined from the sample, and they are allowed to change when states are added to or removed from the sample, changing the sample will change the

distribution used to determine the rounding marks. Determining properties of a distribution from previous censuses while incorporating a model to update those parameters to the current census year is immune to the New State paradox, as long as the parameters are not further updated in response to the inclusion of the new state, but the unbiasedness of such an apportionment is only as accurate as the model.

The Alabama paradox will not occur if rounding marks cannot catch up with and overtake state populations when the Population Ruler is moved. As the population ruler moves to the right, D decreases. Provided that $d \log r / d \log D \geq -1$, rounding marks do not move faster than the population ruler, and the Alabama paradox will not occur.

Taking the derivative of Eq.(7) with respect to D we obtain

$$p(rD)D \frac{drD}{dD} = \frac{1}{D} \int_{fD}^{(f+1)D} p(v)v dv. \quad (9)$$

For this to be true

$$\frac{drD}{dD} \geq 0, \quad (10)$$

which implies

$$\frac{d \log r}{d \log D} \geq -1, \quad (11)$$

and the Alabama paradox can never occur.

We believe that unbiased apportionment methods are the first example in the literature of apportionment methods that avoid the Alabama paradox and the New State paradox, yet are not divisor methods because they are not homogeneous.

In the next section we show that many of the familiar divisor methods are unbiased apportionment methods for power law distributions.

X. POWER LAW DISTRIBUTIONS

Consider the power law population distribution, $p_\beta \propto v^{\beta-1}$. The cumulative distribution function is $I(v) \propto v^\beta$. Solving Eq. (8) for $r_\beta(f)$ we have

$$r_\beta(f) = \frac{1}{D} \left(\frac{1}{D} \int_{fD}^{(f+1)D} v^\beta dv \right)^{1/\beta} = \left(\frac{(f+1)^{\beta+1} - f^{\beta+1}}{\beta+1} \right)^{1/\beta} \quad (12)$$

The rounding mark $r_\beta(f)$ does not depend on D , so power law apportionment methods are homogeneous divisor methods. They are the only unbiased apportionment methods with

TABLE II. Rounding Marks for Power Law Distributions

f	β										
	$-\infty$	-4	-3	-2	-1	0	+1	+2	+3	+4	$+\infty$
0	0.00	0.00	0.00	0.00	0.00	0.37	0.50	0.58	0.63	0.67	1.00
1	1.00	1.36	1.39	1.41	1.44	1.47	1.50	1.53	1.55	1.58	2.00
2	2.00	2.42	2.43	2.45	2.47	2.48	2.50	2.52	2.53	2.55	3.00
3	3.00	3.44	3.45	3.46	3.48	3.49	3.50	3.51	3.52	3.54	4.00
4	4.00	4.45	4.46	4.47	4.48	4.49	4.50	4.51	4.52	4.53	5.00

this property. This is so because power law distribution functions are the only distributions that have the same shape on any interval.

Examples of r_β are shown in Table II. Some values are determined by taking the appropriate limits. Formulas for some of the values are shown below.

$$r_{-\infty}(f) = f, \quad (13)$$

$$r_{-2}(f) = \sqrt{f(f+1)}, \quad (14)$$

$$r_{-1}(f) = (\log(f+1) - \log(f))^{-1}, \quad (15)$$

$$r_0(f) = \frac{(f+1)^{f+1}}{ef^f}, \quad (16)$$

$$r_1(f) = f + \frac{1}{2}, \quad (17)$$

$$r_2(f) = \sqrt{f(f+1) + \frac{1}{3}}, \quad (18)$$

$$r_\infty(f) = f + 1. \quad (19)$$

Several of these rounding rules have special names. $r_{-\infty}$ is the Adams apportionment rounding rule. r_{-2} is the Huntington-Hill rule. r_1 is the Webster rule. r_∞ is The Jefferson rule. There are some divisor methods, such as Dean's method, that are not derivable from any power law.

As Balinski and Young noted,¹ if the states are drawn from a uniform distribution, Webster's method is unbiased. However, if the states are drawn from a different distribution, then Webster's method will be biased. In the next section we determine that no single power law distribution fits the observed distribution of state populations over the entire range of populations.

XI. A MODEL OF POPULATION GROWTH LEADING TO THE LOG-NORMAL DISTRIBUTION

Figure 1 shows histograms of the distributions of the logarithms of state populations for recent censuses¹⁰ at 20-year intervals. The distributions are roughly bell-shaped. There is

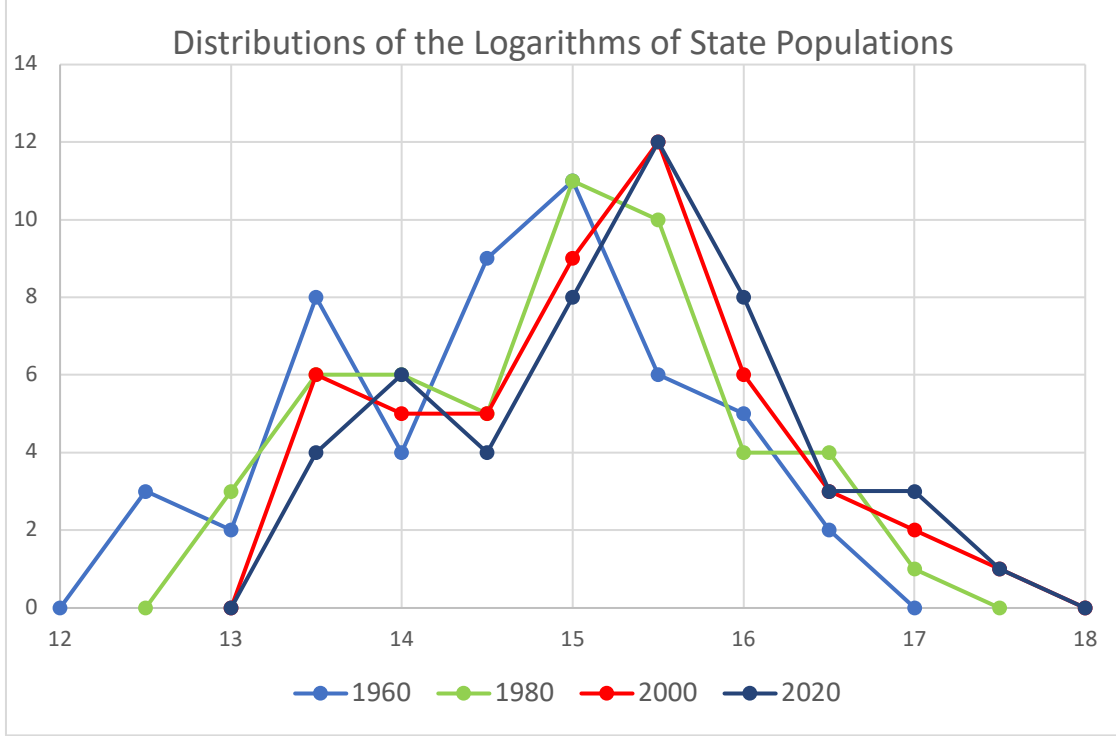


FIG. 1. Histograms of the distributions of the logarithms of state populations are shown for recent censuses at 20-year intervals.

no single power law distribution that can fit these distributions, since a positive exponent is required to fit small populations and a negative exponent is required to fit large populations. The $\beta = 0$ distribution is the power law distribution that fits census data best, in that it maximizes the likelihood of the observed distribution for integer β , but the distribution fails at the tails, which matter the most when considering size bias. Table III shows moments for the logarithm of state populations for recent state population distributions.¹⁰ The skew and excess kurtosis, both zero for normal distributions, are sufficiently small that the log normal distribution is a reasonable approximation for these distributions. The mean of the logarithm of state populations increases almost perfectly linearly with time, with an R^2 value of 0.99. The standard deviation of the logarithm of state populations appears to be

TABLE III. Historical Statistics of $\text{Log}(V)$.

Year	mean	standard deviation	skew	excess kurtosis
2020	15.218	1.024	-0.047	-0.514
2010	15.156	1.019	-0.054	-0.537
2000	15.062	1.020	-0.052	-0.572
1990	14.939	1.018	-0.015	-0.630
1980	14.850	1.021	-0.073	-0.719
1970	14.714	1.063	-0.091	-0.719
1960	14.583	1.071	-0.172	-0.606

decreasing towards a fixed point approximately equal to 1.

We now provide an explanation for the distribution being well approximated by a log normal distribution and for the observed dynamics of the mean and standard deviation of the logarithms of state populations.

A simple model of state population growth is a random rate model,

$$\frac{dv_c(t)}{dt} = g_c v_c(t) + \zeta_c(t) v_c(t), \quad (20)$$

in which g_c is a population growth rate, possibly dependent on population, and $\zeta_c(t)$ is random with zero mean. In terms of the logarithm of a state's population, $x_c(t) = \log(v_c(t))$,

$$x_c(t + \Delta t) = x_c(t) + \Delta t g_c(t) + \Delta \zeta_c(t), \quad (21)$$

where $\langle \Delta \zeta_c(t) \rangle = 0$ and $\langle \Delta \zeta_c(t) \Delta \zeta_d(t') \rangle = \delta_{c,d} \delta_{t,t'} \Delta t 2h$, where h represents the size of fluctuations in the growth rate. The Fokker-Planck equation⁸ determines the dynamics of the distribution governed by the above equations,

$$\partial_t \tilde{p}(x, t) = -\partial_x g(x) \tilde{p}(x, t) + \partial_x^2 h(x) \tilde{p}(x, t). \quad (22)$$

A simple model of state population dynamics takes the form

$$g(x) = g - (x - x_0(t))m \quad (23)$$

$$h(x) = h \quad (24)$$

where $g > 0$ is the average growth rate, $m > 0$ tunes the migration rate from large to small states, and $h > 0$ is the scale of random fluctuations in the growth rate. The formula for

$g(x)$ is determined from the first two terms of the Taylor expansion around the mean of the distribution, and for $m > 0$ it models the effect of migration from large- to small-population states, with a larger growth rate for smaller states and a smaller growth rate for larger states.

Plugging in the normal distribution,

$$\tilde{p}(x, t) = \frac{\exp\left(-\frac{(x-x_0(t))^2}{2\sigma^2(t)}\right)}{\sqrt{2\pi\sigma^2(t)}}, \quad (25)$$

we find that the normal distribution is a solution, with the mean and standard deviation of logarithm of a state's population over time determined by

$$\frac{dx_0}{dt} = g, \quad (26)$$

$$\frac{d\sigma^2}{dt} = 2h - 2m\sigma^2(t). \quad (27)$$

x_0 increases linearly with time, and the standard deviation approaches a stable fixed point $\sigma^2 = h/m$, in agreement with observation. The observed standard deviation appears to be close to 1, implying that m is approximately equal to h . It is unclear if this is just a coincidence or if it arises from something that the model neglects.

Remembering that $x = \log(v)$, the fixed-point distribution of state populations is the log normal distribution,

$$p(v)dv = \tilde{p}(\log v)d\log(v) = \frac{\exp\left(-\frac{\log^2(v/v_g)}{2\sigma^2}\right)}{v\sqrt{2\pi\sigma^2}}dv, \quad (28)$$

where $v_g(t)$ is the geometric mean of the state populations at time t .

XII. UNBIASED APPORTIONMENT FOR STATES DRAWN FROM A LOG NORMAL DISTRIBUTION

We determine the unbiased rounding marks for the log normal distribution from Eq. (8) by integrating and inverting its cumulative distribution function. Expressing the cumulative distribution function in terms of $q = V/D$, we have $I_{LN}(q, \log q_g, \sigma) = I(qD, \log(q_g D), \sigma)$, where $q_g = v_g/D$, $\log q_g$ is the mean of $\log q$, and σ is the standard deviation of $\log q$. The rounding mark for the f family is determined from

$$I_{LN}(r, \log q_g, \sigma) = \int_f^{f+1} I_{LN}(q, \log q_g, \sigma)dq, \quad (29)$$

which equals

$$\begin{aligned}
I_{LN}(r, \log q_g, \sigma) &= I_{LN}((f+1), \log q_g, \sigma)(f+1) \\
&- I_{LN}(f, \log q_g, \sigma)f - I_{LN}((f+1), \sigma^2 + \log q_g, \sigma)e^{\sigma^2/2}q_g \\
&+ I_{LN}(f, \sigma^2 + \log q_g, \sigma)e^{\sigma^2/2}q_g.
\end{aligned} \tag{30}$$

Provided that v_g and σ are determined a priori and not from the actual state populations, the apportionment method associated with these rounding marks is immune to the New State paradox and the Alabama paradox, yet it is not a divisor method. The rounding mark between f and $f+1$ depends on f , as it would for a divisor method, but it also depends on D through q_g , which indicates that it is not homogeneous.

TABLE IV. Rounding Marks for Log Normal Distributions

f	q_g				
	1	2	5	10	20
0	0.491	0.539	0.591	0.623	0.650
1	1.461	1.481	1.506	1.525	1.543
2	2.468	2.480	2.495	2.507	2.518
5	5.479	5.485	5.492	5.497	5.502
10	10.487	10.489	10.493	10.496	10.499
20	20.492	20.493	20.495	20.497	20.498

Table IV shows rounding marks for log normal distributions for various q_g and $\sigma = 1$. Looking at the $q_g = 5$ column, one sees that the fractional part of a family's rounding mark starts above 0.5 and initially decreases with increasing family number, passing through 0.5 near the mode of the distribution, consistent with the idea that as one moves from family to family, from the left end of the distribution to the right end of the distribution, the fractional part of the rounding mark will be on the side of 0.5 where the distribution of states in a family is larger. As the family number continues to increase, the fractional part of the rounding mark continues to fall, eventually increasing again to approach 0.5, just as it does for power law distributions. Looking at the $f = 5$ row, when $q_g = 1$, the 5 family is in the right tail of the distribution so its rounding mark is at the left side of the family. As q_g increases, the rounding marks move to the right as the 5 family moves into the left tail of the distribution.

The log normal apportionment using the mean and standard deviation of the 2020 census populations¹⁰, for which $q_g = 5.34$, does not agree with Webster’s method applied to families. This is because the actual distribution of states has bumps at the tails that deviate from the log normal distribution. The log normal apportionment does agree with Webster applies to states for the 2020 census.

XIII. CONCLUSION

We have presented impartial and unbiased methods for Congressional apportionment. Webster applied to families is an impartial apportionment method that apportions the same number of seats to families of states containing the same number of people regardless of whether a family is composed of a large number of small-population states or a small number of large-population states. Webster applied to families is immune to the Alabama paradox but not the New State paradox. Unbiased methods apportion seats so that if states are drawn repeatedly from the same distribution, the expected value of the total number of seats apportioned to each family equals the expected value of the total divisor-method quota for that family. We found the unbiased apportionment methods for power law and log-normal distributions.

* roshyman@att.net

† ntideman@vt.edu

¹ Balinski, M. and Young H. (2001 [1982]) Fair Representation (2nd. ed.), Brookings Institution. (1st ed.), Yale University Press, New Haven.

² Biles, Charles M. (2015) "The Congressional Apportionment Problem Based on the Census 1790-1840: Basic Divisor Methods.", pp.3-6.

³ Ernst, Lawrence R. (1994) Apportionment Methods for the House of Representatives and the Court Challenges, Management Science, 40 (10), 1207-1227.

⁴ Huntington, E. (1921) A New Method of Apportionment of Representatives, Quart. Publ. Amer. Stat. Assoc, 17, 859-870.

- ⁵ Huntington, E. (1928) The Apportionment of Representatives in Congress, Transactions of the American Mathematical Society, 30, 85-110.
- ⁶ Jefferson, Thomas (1904 [1792]) Opinion on Apportionment Bill, in The Writings of Thomas Jefferson. Washington, DC, vol. 3.
- ⁷ Malkevitch, Joseph (2002) Apportionment, American Mathematical Society, Retrived online 10/19/2021 |<http://www.ams.org/publicoutreach/feature-column/fcarc-apportionii1>.
- ⁸ Pavliotis, Grigorios A. (2014). Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations. Springer Texts in Applied Mathematics. Springer. ISBN 978-1-4939-1322-0.
- ⁹ Pukelsheim, Friedrich (2017 [2014]) Proportional Representation: Apportionment Methods and Their Applications (2nd. Ed.), Springer Nature, Cham, Switzerland.
- ¹⁰ United States Census Bureau (2021) 2020 Census Apportionment Results, Table C2. Apportionment Population and Number of Seats in U.S. House of Representatives by State: 1910 to 2020 Retrieved online 10/19/2021 <https://www.census.gov/data/tables/2020/dec/2020-apportionment-data.html>
<https://www2.census.gov/programs-surveys/decennial/2020/data/apportionment/apportionment-2020-tableC2.xlsx>.