Minimally disproportional representation: generalized entropy
and Stolarsky Mean-Divisor Methods of Apportionment

by

Luc LAUWERS
Tom VAN PUYENBROECK
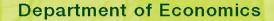
Econometrics

DISCUSSION
PAPER

Minimally Disproportional Representation: Generalized Entropy and Stolarsky Mean-Divisor Methods of Apportionment

Luc Lauwers[†], Tom Van Puyenbroeck[†,‡]

†: Center for Economic Studies, KU Leuven
Naamsestraat 69, 3000 Leuven, Belgium
e-mail: luc.lauwers@econ.kuleuven.be

‡: Research Unit Economics and Policy, HU Brussel
Stormstraat 2, 1000 Brussels, Belgium
e-mail: tom.vanpuyenbroeck@hubrussel.be

July 1, 2008

Abstract

We study divisor methods, the primary class to solve apportionment problems, based upon Stolarsky means $S_{\alpha,\beta}$. These encompass the five traditional methods. We disclose a one-to-one relation between methods of the form $S_{\alpha,1}$ and aggregate measures of seat/vote disproportionality of the generalized entropy family: using a divisor method based upon such a generalized logarithmic mean coincides with minimizing a generalized entropy inequality measure. The Balinski-Young 'favoring small states'-ordering ranks the generalized entropy methods. This framework improves upon an inconsistency in the traditional inequality approach to apportionment problems, which we illustrate by showing that the major rationale of the 'method of equal proportions' is consistently preserved by a non-traditional method.

1 Introduction

Throughout the world, different apportionment methods are used either to allocate seats to parties in proportion to their respective vote totals, or to electoral districts in proportion to their respective population sizes. The 'one man, one vote' tenet summarizes the ideal solution to these fair division problems, but in reality this ideal is only rarely achieved. Since seats are indivisible goods, one normally has to round the ideal, i.e., perfectly proportional, seat numbers to integer values. This rounding explains why in practice all methods entail a residual degree of malapportionment. A second empirical regularity which is relevant to our problem is that these methods, while all aspiring to the ideal of exact proportionality, may well lead to different seat allocations. For example, based on the 2000 census figures, the state of California is theoretically entitled to about 52.45 of the 435 seats in the US House of Congress, and as a result of the 'method of equal proportions' used since 1940 has 53 representatives. Yet even within the class of 'weakly' proportional methods this number could have been as low as 50 or as high as 55, depending on the specific apportionment method used.[1]

If proportional representation aims at "equity as near as may be" (cf. Young (1994), p. 42), and if weakly proportional methods may provide different solutions to the same problem, it must be the case that these methods' underlying fairness conceptions are different. We provide a systematic treatment of this intuition, as we analyse the connection between an entire popular class of apportionment methods and an entire class of inequality measures. Specifically, if minimizing unfairness means that individual voters' seat per vote ratio's should be as equal as possible, our main theorem asserts the following: if inequality so conceived is measured by any member of the generalized entropy class with parameter α, then minimizing voters' inequality is *equivalent* to applying a divisor method of apportionment that uses the Stolarsky mean $S_{\alpha,1}$ with the *same* α-parameter to solve the integer rounding problem. We discuss the broader relevance of this theorem in the next paragraphs.

The idea to couple entropy measures with the apportionment problem is not new. In fact, Henri Theil (1969) introduced his measure in the political science literature immediately after presenting it as an income inequality gauge. Furthermore, Cowell and Kuga (1981, p. 303) concluded one of the seminal articles advocating generalized entropy for income inequality measurement by suggesting that "our methodology may be more attractive in applications [...] outside the Economics discipline —for example in the field of Political Science where one derives so-called 'Indices of Malapportionment'...". With hindsight, one cannot but notice the irony of that fairly defensive statement: the current status of generalized entropy measures as mainstream tools for measuring income inequality starkly contrasts with their very rare application to political representation inequality measurement. But the idea deserves to be enlivened; as the equivalence relation of our main theorem illustrates, generalized entropy

---

[1]By a weakly proportional method we mean a method that would lead to perfect proportionality whenever this is a feasible solution. See section 3 for a formal statement of this property.

measures do occupy an important role in the analysis of proportional representation, extending beyond the convenient (decomposability) properties that they anyhow possess for the positive analysis of malapportionment.

We restrict the scope of our analysis to divisor methods, since these are primary instruments to solve apportionment problems. As pointed out from the late 19th century onwards, such methods prevent inconvenient results that could originate from changes in the variables that constitute an apportionment problem. That is, they avoid the Alabama paradox (occuring if the House size increases and a state/party loses a seat as a result), the population paradox (occuring if one state/party grows at a faster rate than another but has to give in seats to that other state/party), and the new state paradox (occuring when as a result of adding a new state/party to the House the seat number of another state/party increases). Studying their properties at a more fundamental level in many theoretical contributions, Balinski and Young (as e.g. summarized in Balinski and Young, 2001) added to such practical concerns the insight that divisor methods are consistent: any apportionment of seats that is judged to be fair must remain unaltered if one would apply the same method, with the relevant variables, to a subset of the states/parties concerned. Consistency thus embeds a concern for procedural fairness into the fair division problem. As we will point out, consistency also is a necessary condition to connect Stolarsky mean divisor methods with generalized entropy measures.

Given that the parameter α of our main theorem may take any real value, our approach implies that we consider an infinite number of apportionment methods. We thus fill in a significant part of the research agenda set out by Grilli di Cortona *et al.* (1999, p. 87, 100), who called for (i) unveiling the hidden fairness criteria behind any specific apportionment method and (ii) "to design new electoral formulas corresponding to suitable measures of disproportionality".

Furthermore, our analysis bears on the mainstream scholarly debate as it developed over the past century. Following Huntington (e.g., 1921, 1928) the general consensus is that only five so-called traditional divisor methods are really worth considering. As far as the apportionment problem of the US Congress is concerned this set has even been narrowed down further to only two methods, viz., the Hill/Huntington 'method of equal proportions' currently used and Daniel Webster's 'method of major fractions', intellectually restituted as being the least biased divisor method by Balinski and Young in the 1980s. Within this debate, our framework adds two points. First, these traditional methods all build on Stolarsky means. Equity objectives with which they have previously been associated are (cardinally) equivalent to the corresponding member of the generalized entropy class. In this sense, we offer an encompassing framework to analyse the problem of inequality in political representation. Secondly, using generalized entropy measures implies that our assessment of such inequality *fundamentally* differs from Huntington's approach. The traditional methods were derived from inequality tests that check whether a seat transfer between two states may decrease the

difference between these states' respective seat-to-vote ratio's. The inequality measures we put central consider instead the deviations of individual's seat-to-vote ratio's from the overall average, and define representation inequality as a population weighted aggregate of such deviations. Such an approach is logically more coherent than one based on state-wise comparisons of seat-to-vote ratio's. To underscore this we will recast Huntington's original concern in our framework and derive a non-traditional divisor method that coherently accomodates the Fundamental Principle which lead him to forward the method of equal proportions.

Related literature.

The rare appearances of entropy measures in the setting at hand have been partly confined to the problem of assessing the degree of malapportionment (or 'distortion' in the terminology of Schofield, 1982) in a given dataset. Using them as such was justified by refering to their decomposability properties (see Monroe, 1984), or, in the specific case of Theil's original measure, because it is a neutral index, i.e. an index "which [is] not embodied in any of the traditional proportional representation methods" (Pennisi, 1998).[2]

As our equivalence result is all about non-neutrality, it fits instead within the formal, normative tradition that links apportionment methods to their respective equity-maximizing objective functions. This tradition, with a history going back at least to Sainte Laguë (1910), notably includes Huntington's aforementioned analyses as well as his protracted debates with Willcox (1954) about the best method for the US House of Representatives. As far as we know, Theil (1969), Theil and Schrage (1977) and, very recently, Agnew (2008) are the only articles in which an entropy measure appears in this normative fashion. Theil (1969) introduced a proportionality independence requirement —the number of seats assigned to a party i relative to those of another party j should uniquely be determined by these two parties' relative vote ratio— and then connected this prerequisite with entropy as a measure for aggregate disproportionality. Although he remained silent on the rounding problem, we build on the same connection in our generalized analysis. An apparently less well-known article of Theil and Schrage (1977) bears an even closer resemblance to the current paper. As we will point out below, these authors both introduced one specific entropy measure of malapportionment and its associated apportionment method, albeit without identifying the latter as a genuine divisor method.[3] (That method, based on logarithmic-mean rounding, was re-introduced by

---

[2]The search for such a neutral index is valuable in itself, precisely because conventional malapportionment indices implicitly endorse a specific apportionment method, which renders the empirical search for the most proportional apportionment method into a tautological exercise. Evidently, as different apportionment methods are used in reality, this implicit bias carries over to comparative studies that assess the degree of proportionality of, say, various national parliaments.

[3]This method was in fact introduced in one of the appendices to their article. Indeed, a major difference between our paper and Theil's work is that we focus on 'one man, one vote' proportional representation whereas Theil (1969) was primarily concerned with justifying systems in which seats are proportional to either the square or the square root of the number of votes. Subsequent contributions (Theil, 1971; Theil and Schrage, 1977) further reveal his strong normative interest in the square root representation system. Section

Agnew (2008), together with a second one based on the identric mean. Both are limiting cases of our theorem as we will indicate in corrolary 1.) Specifically, Theil and Schrage tackled the rounding issue by embedding their objective of inequality minimization in an integer programming problem. We proceed likewise for the generalized entropy class, and in doing so seek to convey that this is actually a natural family of inequality measures for the apportionment problem.

Ultimately, that claim rests on two well-documented facts which, curiously enough, have not been linked explicitly yet. First, divisor methods constitute an algorithm to rank vote-seat pairs in a priority order and then distribute seats according to this priority untill all available seats are apportioned. This essentially explains the equivalence with solving an integer programming problem with an additively separable objective function. Young (1994, p. 188-189) shows that rank-index methods, which include divisor methods, emerge from a suitably formulated problem of maximizing a sum of 'priorities'. Balinski and Ramirez (1999) consider a family of additive functions that measure the deviation of a state's seat number to its ideal, and link them with stationary signpost divisor methods.[4] Grilli di Cortona *et al.* (1999, p. 93-97) discuss voter-oriented inequality minimization problems with additively separable objective functions, and point at the strong resemblance between their 'greedy algorithms' and (the traditional) divisor methods (our Proposition 2 identifies an additional condition for their equivalence). A second fact was, however, overlooked in previous analyses, although it is a central result in the income inequality literature: if one wants this objective function to be a regular relative inequality measure, then additive separablity narrows down the spectrum of eligible functions to the generalized entropy class (see e.g. Chakravarty, 1999).[5] In this sense, our central claim is both obvious and compelling.

The rest of this paper unfolds as follows. Section 2 introduces notation, recalls some essential features of divisor methods and the definition of Stolarsky means, and investigates within the class of Stolarsky divisor methods the relation "favoring small states". Section 3 addresses the issues sketched in this Introduction in a formal framework and provides our main theorem. Section 4 re-examines the traditional inequality analysis of divisor methods, and provides a non-traditional alternative for the method of equal proportions. Section 5 gathers some additional comments. Most of the proofs are collected in an appendix.

---

5 returns to this observation.

[4] In the notation of section 2, stationary signpost divisor methods use signposts $d(k) = k + c$, where c is a constant such that $0 \leq c \leq 1$. This class contains three of the five traditional methods, viz. Jefferson, Webster, and Adams.

[5] Following Chakravarty (1999), we call a measure a regular inequality measure when it is (a) anonymous, (b) obeys Dalton's transfer principle, (c) obeys the Population principle, and (d) is normalized such that perfect equality yields a zero value for the inequality measure. A *relative* inequality measure is such that its value is unaffected by equiproportional changes in the equalisandum. In a proportional representation-context this is a natural requirement.

2 Divisor methods and Stolarsky means

A specific problem of apportionment is an $(n + 1)$-tuple (p, H) with $p = (p_1, p_2, \ldots, p_n)$ an n-tuple of populations (e.g. counts of population in n states, counts of votes for n parties, ...) and H an integer size of the house. Let $\mathcal{P}^n$ collect all apportionment problems. Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ collect the natural numbers. An apportionment method M is a correspondence

$$M : \mathcal{P}^n \longrightarrow\!\!\!\rightarrow \mathbb{N}^n : (p, H) \longmapsto M(p, H),$$

enforcing the condition $a_1 + a_2 + \cdots + a_n = H$ for each a in $M(p, H)$.

The proportionality requirement invokes the need for specifying desirable properties of M. For the literature on fair apportionment the monograph by Balinski and Young (2001) is the basic reference. As stated in the Introduction, one of their major conclusions is that divisor methods form the primary class of apportionment methods. Such a divisor method is defined through a sequence $d : \mathbb{N} \to \mathbb{R} : k \mapsto d(k)$ of signposts. The signpost $d(k)$ belongs to the closed interval $[k, k + 1]$ and splits the interval into a left part where numbers are rounded down to k, and a right part where numbers are rounded up to $k + 1$. For $d(k)$ itself, there is the option to round down to k or to round up to $k+1$, thus possibly generating multiplicities. In other words, the d-rounded number $[y]_d$ with y in $\mathbb{R}^+$ is equal to k in $\mathbb{N}$, in case

$$k \leq y \leq d(k) \leq k + 1 \qquad \text{or} \qquad k - 1 \leq d(k - 1) \leq y \leq k.$$

The divisor method based upon the sequence d —denoted D_d— solves the problem (p, H) by looking for a common divisor x that forces the rounded quotients $[p_i/x]_d$ to add up to the number H. Equivalently, a belongs to $D_d(p, H)$ if and only if there exists an x in $\mathbb{R}^+$ such that for each i we have either "$a_i > 0$ and $p_i/d(a_i - 1) \geq x \geq p_i/d(a_i)$", or "$a_i = 0$ and $x \geq p_i/d(a_i)$". This leads to the following Balinski-Young characterization:

$$D_d(p, h) = \left\{ a : \min_{a_i > 0} \frac{p_i}{d(a_i - 1)} \geq \max_{a_j \geq 0} \frac{p_j}{d(a_j)} \quad and \quad \sum_i^n a_i = H \right\}, \tag{1}$$

where one assumes that $p_i > p_j$ implies $p_i/0 > p_j/0 > p_i$. The characterization (1) of divisor methods conveys that the transfer of the a_i-th seat from i to j is unfair in view of the priority standard $p/d(a_i)$. Thus, party i gets exactly a_i seats because it 'deserves' the a_i-th seat before party j deserves its $(a_j + 1)$-th seat (for each $j \neq i$). The priority standard also constitutes an algorithmic representation of the divisor method D_d: starting from $D_d(p, 0) = (0, 0, ..., 0)$, each transition from $D_d(p, H')$ to $D_d(p, H' + 1)$ involves giving that seat to party k with $p_k/d(a_k) = \max_i p_i/d(a_i)$, untill the H seats have been allocated.

Five methods belonging to this class are often grouped together as the 'five traditional methods'. The Jefferson method rounds down all quotients, hence employing $d(k) = k + 1$. The Webster method employs standard rounding with the arithmetic mean of k and $k + 1$

defining the dividing point. The rounding-up rule $d(k) = k$ is associated with the Adams method. The two other traditional methods are those of Dean and the Hill/Huntington method of Equal Proportions, which respectively use the harmonic mean and the geometric mean to split the closed interval $[k, k+1]$-interval. A divisor method for which $d(0) = 0$, as e.g., Adams, Dean, and Hill, allocates at least one seat to each party if $n \leq H$, otherwise (if $H < n$) the seats are allocated to the H largest populations.

The important position of these five methods in many apportionment debates notwithstanding, it is evident from the foregoing that the class of divisor methods is infinitely large. Given this abundance one is invited to bring in some structure. Balinski and Young proposed to compare different methods according to whether they favor small states, and obtain the next result.

Theorem 1 (Balinski and Young, 2001, Thm 5.1)**.** Let $D_{d'}$ and D_d be two divisor methods. Let $d'(a)/d'(b) > d(a)/d(b)$ for all a and b in $\mathbb{N}$ with $a > b$. Then, $D_{d'}$ favors small states relative to D_d, that is, for each apportionment problem (p, H) in $\mathcal{P}^n$, for each a in $D_d(p, H)$, and for each a' in $D_{d'}(p, H)$, we have

$$p_i < p_j \quad \text{implies either} \quad a_i' \geq a_i \quad \text{or} \quad a_j' \leq a_j \,.$$

Equivalently, using the terminology of Marshall, Olkin, and Pükelsheim (2002), Theorem 1 identifies the condition for the fact that method $D_{d'}$ *"is majorized by"* method D_d.

We now turn to divisor methods based upon Stolarsky means (named after Kenneth B. Stolarsky, 1975):

$$S_{\alpha,\beta} : \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+ : (x, y) \longmapsto \left(\frac{\beta}{\alpha} \times \frac{x^\alpha - y^\alpha}{x^\beta - y^\beta} \right)^{1/(\alpha - \beta)}, \tag{2}$$

with α and β in $\mathbb{R}$. These means were introduced for $\beta \neq \alpha \neq 0$ and continuously extended to all pairs α and β in $\mathbb{R}$. The values

$$S_{\alpha,0}(x, y) = \left(\frac{x^\alpha - y^\alpha}{\alpha (\ln x - \ln y)} \right)^{1/\alpha} \quad \text{and} \quad S_{\alpha,\alpha}(x, y) = \exp\left(-\frac{1}{\alpha} + \frac{x^\alpha \ln x - y^\alpha \ln y}{x^\alpha - y^\alpha} \right)$$

are known as the logarithmic and the identric mean of order α, respectively.

The value $S_{\alpha,\beta}(x, y)$ fits the definition of a mean. Stolarsky (1975) pointed out that $\lim_{\alpha \to 0} S_{\alpha,1}(x, y)$ is the logarithmic mean, and introduced the identric mean as $\lim_{\alpha \to 1} S_{\alpha,1}(x, y)$. Importantly, this family also comprises the means that underlie the traditional divisor methods. Indeed, consider the divisor sequence

$$d_{\alpha,\beta} : \mathbb{N} \longrightarrow \mathbb{R} : k \longmapsto S_{\alpha,\beta}(k+1, k),$$

and denote by $D_{\alpha,\beta}$ the corresponding divisor method. Table 1 represents the five traditional divisor methods (Adams, Dean, Hill, Webster, and Jefferson) in terms of Stolarsky methods. For later reference, the methods $D_{0,1}, D_{\frac{1}{2},1}$ and $D_{1,1}$ are also included.

Within the class of Stolarsky divisor methods the following relationships hold.

Lemma 1. Let $\alpha \geq \alpha'$. Then, the divisor method $D_{\alpha',1}$ favors small states relative to the divisor method $D_{\alpha,1}$.

Proof. If $\alpha = \alpha'$, then $D_{\alpha,1} = D_{\alpha',1}$. Hence, let $\alpha > \alpha'$. We show that the sequence $k \mapsto f(k) = S_{\alpha',1}(k+1,k)/S_{\alpha,1}(k+1,k)$ is strictly increasing in k. Apply the logarithm to

$$f(k) = \frac{\left[\frac{(k+1)^{\alpha'} - k^{\alpha'}}{\alpha'}\right]^{1/(\alpha'-1)}}{\left[\frac{(k+1)^{\alpha} - k^{\alpha}}{\alpha}\right]^{1/(\alpha-1)}}$$

and differentiate with respect to k (we extend the domain of f to $\mathbb{R}^+$). Conclude that f is strictly increasing if

$$\frac{\alpha'}{\alpha' - 1} \frac{(k+1)^{\alpha'-1} - k^{\alpha'-1}}{(k+1)^{\alpha'} - k^{\alpha'}} - \frac{\alpha}{\alpha - 1} \frac{(k+1)^{\alpha-1} - k^{\alpha-1}}{(k+1)^{\alpha} - k^{\alpha}} > 0.$$

This inequality holds if

$$\frac{\alpha'}{\alpha} \frac{(k+1)^{\alpha} - k^{\alpha}}{(k+1)^{\alpha'} - k^{\alpha'}} > \frac{\alpha' - 1}{\alpha - 1} \frac{(k+1)^{\alpha-1} - k^{\alpha-1}}{(k+1)^{\alpha'-1} - k^{\alpha'-1}}.$$

If $\alpha > \alpha'$, then the map $x \mapsto x^{1/(\alpha-\alpha')}$ is strictly increasing, and the previous inequality is fulfilled if and only if

$$S_{\alpha,\alpha'}(k+1,k) > S_{\alpha-1,\alpha'-1}(k+1,k).$$

Since $S_{\alpha,\beta}(x,y)$ is strictly increasing in both α and β (Stolarsky, 1975), the latter inequality indeed holds and the result follows. $\square$

Lemma 1 is restricted to divisor methods based upon Stolarsky means $S_{\alpha,1}$. These will be assigned a prominent role in the next sections. However, we point out that the result in Lemma 1 extends.

Proposition 1. Let $\alpha \geq \alpha'$ and $\beta \geq \beta'$. Then,

- the method $D_{\alpha',\beta}$ favors small states relative to the method $D_{\alpha,\beta}$ (where $\beta \neq 0$),
- the method $D_{\alpha,\beta'}$ favors small states relative to the method $D_{\alpha,\beta}$ (where $\alpha \neq 0$).

Proof. See Appendix.

Marshall et al (2002, Proposition 2) order divisor methods based upon the signpost sequence $S_{2\alpha,\alpha}(k+1,k)$ of power means. Proposition 1 accommodates their result. Furthermore,

the power mean $S_{2\alpha,\alpha}(a+1, a)$ converges to a as α goes to $-\infty$. Hence, Table 1 provides the following extension of the traditional sequence

$$\text{Adams} > \text{Dean} > \text{Hill} > D_{0,1} > D_{\frac{1}{2},1} > D_{1,1} > \text{Webster} > \text{Jefferson},$$

where the symbol ">" stands for "favors small states relative to". Table 2 illustrates the position of the methods $D_{0,1}$, $D_{\frac{1}{2},1}$, and $D_{1,1}$ as compared to the five traditional methods, for a problem in which $n = 10$, $\sum_i^n p_i = 100\,725$, and $H = 100$.

3 Generalized entropy and Stolarsky divisor methods

We now recast the central issues of this paper in a formal framework. Consider an apportionment problem (p, H) and let $p_1 + p_2 + \cdots + p_n = P$. The "ideal allocation" is described by

$$q = (p_1\,H/P,\ p_2\,H/P,\ \ldots,\ p_n\,H/P)\,.$$

The value $q_i = p_i H/P$ is customarily referred to as the *quota* of state i. Let M be an apportionment method. It is desirable that each allocation a in $M(p, H)$ is close to the allocation q. In particular, the method should be *weakly proportional*: if q happens to be an n-tuple of natural numbers, then $M(p, H) = \{q\}$ and the n-tuples

$$a/H = (a_1/H,\ a_2/H,\ \ldots, a_n/H) \qquad \text{and} \qquad p/P = (p_1/P,\ p_2/P,\ \ldots, p_n/P)$$

coincide.[6] In most cases, however, the constraints $a_i \in \mathbb{N}$ force a gap between a/H and p/P. We propose to use a map $\Phi : \mathbb{Q}^n \times \mathbb{Q}_0^n \to \mathbb{R}$ to measure the inequality between a/H and p/P and to select inequality minimizing allocations. Obviously, for each y in $\mathbb{Q}^n$ the value $\Phi(x, y)$ is minimal for $x = y$.[7] In particular, the value $\Phi(x, p/P)$ is minimal for $x = q/H$.

Let us first spell out the procedural consequences of imposing additional structure on M and Φ. We require that the method M is *anonymous*[8] and that the measure Φ is *additively separable*:

$$\Phi(a/H,\ p/P) = \varphi(a_1/H,\ p_1/P) + \varphi(a_2/H,\ p_2/P) + \cdots + \varphi(a_n/H,\ p_n/P),$$

---

[6]This notion of weak proportionality follows Balinski and Young (2001). A different version, proposed by Theil (1969), will be labelled "proportionality independence".

[7]A shift in the map Φ might result in $\Phi(x, x) = 0$.

[8]Hence, permuting the populations of the states permutes the allocated seats in the same way. As a consequence, the measure Φ should be anonymous in the qualified sense of Cowell (1980): the couples (a_i, p_i) $i = 1, 2, ..., n$ may be re-ordered without changing the total value of the measure, but the components of a and p cannot be permuted independently. Measure (4) lives up to this requirement.

Table 1: Eight Stolarsky Mean Divisor Methods

Divisor Method	$d(a)$	$S_{\alpha,\beta}$	
Adams	a	$\alpha \to -\infty$	$\beta = 1$
Dean	$a(a+1)/(a+0.5)$	$\alpha = -2$	$\beta = -1$
Hill	$\sqrt{a(a+1)}$	$\alpha = -1$	$\beta = 1$
$D_{0,1}$	$\frac{1}{\ln(a+1)-\ln a}$	$\alpha = 0$	$\beta = 1$
$D_{\frac{1}{2},1}$	$\left(\frac{\sqrt{a+1}-\sqrt{a}}{1/2}\right)^{-2}$	$\alpha = 1/2$	$\beta = 1$
$D_{1,1}$	$\left(\frac{1}{e}\frac{(a+1)^{a+1}}{a^a}\right)^{-1}$	$\alpha = 1$	$\beta = 1$
Webster	$a + \frac{1}{2}$	$\alpha = 2$	$\beta = 1$
Jefferson	$a + 1$	$\alpha \to \infty$	$\beta = 1$

Table 2: Stolarsky Divisor Methods in the 'favoring small states'-ordering

Voters	(ideal)	Adams	Dean	Hill	$D_{0,1}$	$D_{\frac{1}{2},1}$	$D_{1,1}$	Webster	Jefferson
42,659	(42.35)	40	42	42	42	42	43	43	44
13,048	(12.95)	13	13	13	13	13	13	13	13
10,879	(10.80)	11	11	11	11	11	11	11	11
10,535	(10.46)	10	10	10	10	11	10	11	10
9,700	(9.63)	10	9	10	10	10	10	10	10
6,500	(6.45)	7	6	6	6	6	6	6	6
2,502	(2.48)	3	3	2	3	3	3	2	2
1,991	(1.98)	2	2	2	2	2	2	2	2
1,461	(1.45)	2	2	2	2	1	1	1	1
1,450	(1.44)	2	2	2	1	1	1	1	1
100,725		100	100	100	100	100	100	100	100

with φ a map from $\mathbb{Q} \times \mathbb{Q}_0$ to $\mathbb{R}$. Furthermore, in order to focus on relative inequality, we require M to be *homogeneous*: $M(p, H) = M(\lambda p, H)$ for each λ in $\mathbb{R}^{++}$. In terms of Φ we require that, for each t and s in $\mathbb{N}$ and each x and y in $\mathbb{Q}^n$, the value $\Phi(t\, x, s\, y)$ is a positive affine transformation of $\Phi(x, y)$, that is, $\Phi(t\, x, s\, y) = \alpha(t, s)\, \Phi(x, y) + \beta(t, s)$ for some $\alpha(t, s)$ in $\mathbb{R}^{++}$ and $\beta(t, s)$ in $\mathbb{R}$. As a consequence, the notation slightly simplifies. In sum, we will solve the problem (p, H) through the next optimization exercise:

$$
\begin{array}{|l}
\text{minimize } \Phi(a, p) \; = \; \varphi(a_1,\, p_1) + \varphi(a_2,\, p_2) + \cdots + \varphi(a_n,\, p_n), \\[2mm]
\text{with respect to } a_1, a_2, \ldots, a_n \text{ in } \mathbb{N}, \text{ and} \qquad\qquad\qquad\qquad\text{(P)} \\[2mm]
\text{subject to } a_1 + a_2 + \cdots + a_n = H.
\end{array}
$$

We use the term *disproportionality method* and the notation I_φ as a reference for the apportionment method obtained by minimizing inequality as described by the optimization problem (P). If the house happens to be less than the number of states, the optimal allocation provides one seat for the larger states. In view of this requirement the assumption (or convention) that $\varphi(1, 0) - \varphi(0, q)$ is strictly decreasing in q is necessary.

The additive separability of Φ is crucial. When measuring income inequality this functional form ensures that the effect on total inequality of a transfer between two individuals depends on (the distance between) their respective income levels only (e.g. Cowell and Kuga, 1981). In the malapportionment framework it likewise implies that one can isolate the effect of a seat transfer between two states. It also implies that the ideal solutions themselves can be considered at a disaggregate level, as notably done by Theil (1969) when introducing the proportionality independence condition $(a_i/a_j) = (p_i/p_j)$ *for all* $i, j = 1, 2, ..., n$. Put otherwise, using an additive separable disproportionality measure Φ forces the method I_φ to be *consistent* (i.e., 'uniform' as defined by Balinski and Young, 2001, p. 141). Consistency as applied upon the states i and j precisely asserts that (a_i, a_j) belongs to $I_\varphi(p_i, p_j; a_i + a_j)$ in case a belongs to $I_\varphi(p, H)$.

Divisor methods do satisfy consistency. On the other hand, the class of consistent methods is larger than the class of divisor methods. Essentially, divisor methods satisfy *weak population monotonicity*: $p_i > p_j$ implies $a_i \geq a_j$. Referring to Theorem 8.4 in Balinski and Young (2001), this property completes the list of five necessary conditions for a divisor method (as we have already incorporated symmetry, weak proportionality, homogeneity and consistency). The next proposition lists further restrictions on φ in order to guarantee that I_φ is a divisor method.

Proposition 2. Let I_φ be the disproportionality method associated with the objective φ. Let the difference

$$
\Delta\varphi(b, q) \; = \; \varphi(b + 1, q) - \varphi(b, q)
$$

be strictly decreasing in q, for each b in $\mathbb{N}$; and strictly increasing in b, for each q in $\mathbb{N}_0$. Then, the disproportionality method I_φ is a divisor method.

Proof. See Appendix.

The differences $\Delta\varphi(b,q)$ allow to decompose the value $\varphi(b,q)$:

$$\varphi(b,q) = \varphi(0,q) + \Delta\varphi(0,q) + \Delta\varphi(1,q) + \Delta\varphi(2,q) + \cdots + \Delta\varphi(b-1,q).$$

The following recursively defined apportionment algorithm solves the integer optimization problem (P):

- $I_\varphi(p,0) = (0,0,\ldots,0)$[9]

- if $a \in I_\varphi(p,H)$ and $\Delta\varphi(a_i,p_i) \leq \Delta\varphi(a_j,p_j)$ for each j, and $e_i = (0,0,\ldots,\overset{\underset{\downarrow}{i}}{1},\ldots,0)$; then $a+e_i$ belongs to $I_\varphi(p,H+1)$.

In words, in allocating the seats one by one, the party with the lowest *marginal disproportionality* receives the next seat. The next equilibrium condition translates this algorithm. The allocation a solves problem (P) if and only if

$$\max_{i|a_i>0}\left\{\Delta\varphi(a_i-1,p_i)\right\} \leq \min_{j}\left\{\Delta\varphi(a_j,p_j)\right\}. \tag{3}$$

Indeed, the inequality $\Delta\varphi(a_i-1,p_i) > \Delta\varphi(a_j,p_j)$ occurs for some i and j if and only if a shift of one seat from i towards j lowers the disproportionality. The restriction that $\Delta\varphi(b,q)$ is decreasing in q and increasing in b guarantees that the optimal allocations are stable against a move of two (or more) seats. The equilibrium described in (3) is similar to the rank-index equilibrium of Balinski and Young (2001, p. 142). According to a rank-index method, the party with the highest value of the rank-index has the priority of receiving the next seat. As a consequence, the results of Balinski and Young (2001, Thm 8.3, Corr 8.3.1) on rank-index methods can be translated: the disproportionality method is consistent and avoids the Alabama paradox. Grilli di Cortona *et al.* (1999, Lemma 6.3, Thm 6.1) obtain similar results using a disproportionality measure of the form $\Theta(a) = \theta(a_1) + \theta(a_2) + \cdots + \theta(a_n)$.

We now present a particular class of regular relative disproportionality measures, viz. those in which the objective Φ is taken from the generalized entropy family. Consider hence

$$J_\alpha : \mathbb{N}^n \times \mathbb{N}_0^n \longrightarrow \mathbb{R} : (a,p) \longmapsto \frac{1}{\alpha(\alpha-1)} \sum_{i=1}^{n}\left[\left(\frac{a_i}{H}\right)^\alpha \left(\frac{p_i}{P}\right)^{1-\alpha} - 1\right], \tag{4}$$

with α a real number different from 0 and 1, $\sum p_i = P$, and $\sum a_i = H$. The value $J_\alpha(a,p)$ can be rewritten as

$$J_\alpha(a,p) = \frac{1}{n\alpha(\alpha-1)} \sum_{i=1}^{n} \frac{p_i}{\mu_p}\left[\left(\frac{a_i/\mu_a}{p_i/\mu_p}\right)^\alpha - 1\right], \tag{5}$$

---

[9]In case $\varphi(0,q) = \infty$ and $H \geq n$, then the algorithm should start with "$I_\varphi(p,n) = (1,1,\ldots,1)$". We refer to the appendix for details.

with μ_a the average seat share H/n, and μ_p the average population share P/n. The objective functions J_α have been given different interpretations. Cowell (1980) introduced them as measures of distributional change, a description that is adequate in the current context as well: they capture the average degree by which ('prior'/'reference') vote shares differ from their respective ('posterior'/'actual') seat shares. Jenkins and O'Higgins (1989) used them to measure income inequality with 'norm incomes' serving as the reference point. The role of the norm income is here assumed by the overall average H/P. As such, they do have a straighforward interpretation in terms of aggregate malapportionment inequality, measuring how much seat-to-vote ratio's a_i/p_i differ from the equality benchmark H/P.

Let us consider problem (P) with $\Phi = J_\alpha$. The corresponding "φ-map" is a positive affine transform of

$$\iota_\alpha : \mathbb{N} \times \mathbb{N}_0 \longrightarrow \mathbb{R} : (b,q) \longmapsto b^\alpha q^{1-\alpha} / \alpha(\alpha - 1).$$

For each b in $\mathbb{N}$, the sign of $(b+1)^\gamma - b^\gamma$ is well defined and equal to the sign of γ. Therefore, the difference

$$\Delta_\alpha(b,q) \;=\; \Delta\iota_\alpha(b,q) \;=\; [(b+1)^\alpha - b^\alpha]\, q^{1-\alpha}/\alpha(\alpha - 1) \tag{6}$$

is strictly decreasing in q and strictly increasing in b. In conclusion, if $\alpha \notin \{0,1\}$, then the disproportionality method $I_\alpha = I_{\iota_\alpha}$ is a divisor method. Theorem 2 investigates the divisor sequence d_α behind I_α.

Theorem 2. Let α be in $\mathbb{R}$, let J_α be the entropy measure as defined by (4), and let $S_{\alpha,1}$ refer to the Stolarsky mean. Then, the disproportionality method I_α and the divisor method $D_{\alpha,1}$ coincide.

Proof. See appendix.

The limiting behavior $\alpha \to 0$ and $\alpha \to 1$ of (4) is well documented. These are interesting insofar they refer to the two original entropy measures as introduced by Theil. Furthermore, I_0 and I_1 satisfy the conditions of Proposition 2 and are divisor methods. We highlight the corresponding results in a separate corollary.

Corollary 1. The divisor method I_0 uses the *logarithmic mean* $S_{0,1}(a+1,a)$ as signpost sequence to minimize proportional representation inequality as measured by

$$J_0(a,p) = \sum_{i=1}^n \frac{p_i}{P} \ln\left(\frac{p_i}{a_i} \frac{H}{P} \right). \tag{7}$$

The divisor method I_1 uses the *identric mean* $S_{1,1}(a+1,a)$ as signpost sequence to minimize proportional representation inequality as measured by

$$J_1(a,p) = \sum_{i=1}^n \frac{a_i}{H} \ln\left(\frac{a_i}{p_i} \frac{P}{H} \right). \tag{8}$$

Proof. See appendix.

While both 'non-traditional', I_0 and I_1 have been considered earlier. Theil (1969) discussed (8) without however addressing the integer rounding problem. On the other hand, Theil and Schrage (1977, p. 262) did solve the problem (P) for (7) and showed that the corresponding seat distribution algorithm hinges on consecutive application of the priority rule

$$\max_i p_i \left(\ln(a_i + 1) - \ln(a_i) \right),$$

which can be recognized as a limiting case of the more general rule (6). Here we complement their original formulation by emphasising that it constitutes the genuine divisor method I_0.

Recently, Agnew (2008) also considered both logarithmic mean rounding and identric mean rounding divisor methods, and linked them respectively to the objective functions $\sum q_i (\ln q_i - \ln a_i)$ and $\sum a_i (\ln a_i - \ln q_i)$.[10] Recalling that $q_i = p_i H/P$, both are easily recognized as (7) resp. (8) up to a multiplicative constant. Conceptually, of course, Agnew's objective functions are geared towards measuring the weighted deviation of seats from their ideal quota. In fact, an appealing feature of Theorem 2 is that it encompasses many previous specific ('case-by-case') results regarding the link between additively separable disproportionality measures and methods of apportionment. For instance, Sainte Laguë (1910) originally retrieved Webster's rule as the solution to the least square problem $\min_{a_i} \Sigma p_i (\frac{a_i}{p_i} - \frac{H}{P})^2$. The latter function is a positive affine transformation of $J_2(a,p)$. Similarly, Grilli di Cortona *et al.* (1999) associate the method of equal proportions with the objective function $\Sigma p_i (\frac{p_i}{a_i} - \frac{P}{H})$, which boils down to $J_{-1}(a,p)$. The fact that Jefferson's method minimizes $\max_i a_i/p_i$ is also consistent with the interpretation of $J_{+\infty}$, etc.

To end this section, recall from table 1 that the method of Dean is based upon the harmonic mean $S_{-2,-1}(a, a+1)$. Hence it is not covered by Theorem 2, and therefore there is no member of the generalized entropy class (4) behind Dean's method. It is worthwhile to indicate the intuition behind this negative result. As pointed out by Cowell (1980), measure (4) [or (5)] assigns the role of (prior) reference distribution to the p_i's and that of the posterior distribution to the a_i's. However, Dean's method intrinsically follows a reverse logic, given that it seeks to equivalise the number of voters per representative. In that sense, it is not surprising that Dean's method does minimize a 'reverse' objective akin to (5), viz.

$$\tilde{J}_2(p,a) = \frac{1}{2n} \sum_{i=1}^{n} \frac{a_i}{\mu_a} \left[\left(\frac{p_i/\mu_p}{a_i/\mu_a} \right)^2 - 1 \right].$$

The question whether the n-tuple a or p should be chosen as the reference distribution returns in the next section.

---

[10]On the connection between (7) and the objective function $\sum q_i (\ln q_i - \ln a_i)$, see also Te Riele (1978).

4 A revision of Huntington's approach

The prominent status of the five traditional divisor methods is largely due to Huntington, who recognized that any apportionment as a rule violates the ideal $a_i/p_i = a_j/p_j$ for all states i, j, and then focused on the quintessential idea that

> "in any practical case, the inequality between these two numbers —that is, the inequality between the two congressional districts A/a and B/b [a_i/p_i and a_j/p_j in our notation]— may be taken as a measure of *the 'amount of inequality' between the two states A and B*. If this inequality can be reduced by a transfer of a representative from one state to the other, then, according to this first criterion, the transfer should be made. The rather vague concept of the inequality *between two states* is thus reduced to the more definite concept of the inequality between two numbers. The question then comes down to this: what shall be meant by the inequality between these two numbers?" (Huntington, 1928, p. 86; italics added)

To illustrate his core idea, assume that there is a pair of states i, j such that $a_i/p_i > a_j/p_j$. One can then check whether $a_i/p_i - a_j/p_j \gtrless (a_j+1)/p_j - (a_i-1)/p_i$, and transfer a seat from i to j if the righthand side turns out to be smaller. Huntington called an apportionment method *workable* if it is always possible to arrive at a final apportionment by successively applying such a pairwise test. Within this framework two problems remain. First, why focus on the number of seats per voter a_i/p_i and a_j/p_j? Huntington in fact also considered a second criterion, based on the fractions p_i/a_i and p_j/a_j (the "number of people per representative"; "the size of a congressional district"). Second, there is the problem of the exact specification of the pairwise inequality tests. These can be based on absolute differences as above (e.g. $a_i/p_i - a_j/p_j$, or $p_i/a_i - p_j/a_j$ if one looks at the second criterion), on relative differences $(((a_i/p_i)/(a_j/p_j)) - 1)$, and the various ways in which such differences can be rewritten. Huntington's (1928) major contribution was to prove that only the five traditional divisor methods emerge as workable. Moreover, the exact specification of the pairwise inequality test drives the choice between these methods; using $a_i/p_i - a_j/p_j$ implies Webster's method, $p_i/a_i - p_j/a_j$ implies Dean's method, $((a_i/p_i)/(a_j/p_j)) - 1$ implies the method of equal proportions, etc. See e.g. Balinski and Young (2001, p. 102) for a further discussion of this issue.

The above approach served as the analytical background for much of the debates on the most appropriate apportionment method for the US House of Representatives. One feature has apparently been overlooked in these discussions, although it constitutes the achilles heel of the approach. Both the above quote and the pairwise inequality tests make clear that a Huntington type of analysis compares seat-to-vote ratio's (or their inverses) *between states*. The number of voters appears in the ratio to be equivalised, but the state-wise inequality comparisons subsequently neglect that different numbers of voters are influenced by a seat

transfer between their respective states. Precisely in this respect, the generalized entropy inequality framework is fundamentally different. A measure like (5) basically works in two stages. First, if the seats per voter in state i deviate from the ideal, then $\left[\left(\frac{a_i/\mu_a}{p_i/\mu_p}\right)^{\alpha} - 1\right]$ gets a non-zero value. Second, an aggregate value of malapportionment is obtained *by weighting* (and summing) these state-specific deviations *by their voter share*. Thus, while the traditional methods also emerge from our framework —together with many other methods that are all 'workable' in the alternative sense of being compatible with Theorem 2— the nature of their constituent inequality tests in the generalized entropy framework is voter-oriented rather than state-oriented (see appendix B for a further discussion of this issue).

Importantly, choosing between these orientations is not a mere question of preference between two possibly relevant units of account. In fact, if the premise is that seat–to-vote ratio's should be as equal as possible, then comparing equally weighted seat-to-vote ratio's fits uneasily with the benchmark role of H/P. Since $\frac{1}{n}\sum_{i=1}^{n}\frac{a_i}{p_i} \neq \frac{H}{P}$, fair (re-)distribution of seats requires taking into account the appropriate relative frequencies in order to restitute the desired equality between the benchmark and the average seat-to-vote ratio. A voter-oriented concern thus logically follows from the choice to equivalise these ratio's, which in turn instigates to obey the elementary 'bookkeeping' equation

$$\sum_{i=1}^{n}\frac{p_i}{P}\frac{a_i}{p_i} = \frac{H}{P}. \tag{9}$$

Intuitively, if one takes it that the equality target is H/P, then the recipients of ultimate relevance logically are the P individuals. Conversely, the problem of dividing H seats over P individuals is indeed one of distributing seat-to-vote ratio's, a result that again directly follows from (9) (on this issue, see also Van Puyenbroeck, forthc.). In that sense, taking "the inequality between the two congressional districts [a_i/p_i and a_j/p_j; hence, seats *per voter*] as a measure of the 'amount of inequality' between the two *states*" is a non-starter.

Let us use our framework to re-address the additional reasons leading Huntington to single out the method of equal proportions as the best possible method. His case rested on two points, viz. (*i*) that this method was not systematically biased in favor of either the smaller or the larger states, and (*ii*) that relative differences are more appropriate than absolute differences to measure pairwise inequality.[11]   Part of Balinski and Young's work

---

[11]The interpretation of that quality has been somewhat dubious. In the main text we connect it with the reference-distribution neutrality property of relative differences, as expressed by the Fundamental Principle. A second interpretation, that relative differences are somehow intrinsically better, is far more controversial. Huntington and others sometimes offered misleading justifications for it, by making such claims as "if the size of the congressional districts is large, say 250,000 in one state and 250,005 in the other, then the difference of five people is of little consequence in so large a number. But if the districts were themselves very small, say 10 and 15, then the same difference becomes important; 15, we say, is larger than 10 by fifty per cent, while 250,005 is larger than 250,000 by only (1/500)th of 1 per cent." (Huntington, 1928, p. 86) Such examples

can be taken as a falsification of the first claim; it is indeed generally acknowledged today that the method of equal proportions has a built-in bias towards the smaller states, and that Webster's method is unbiased towards states. His second point rested on a symmetry consideration, as is apparent from his *Fundamental Principle* that (in our notation) "in a satisfactory apportionment between two states i and j, we shall agree that a_i/p_i and a_j/p_j should be as near as possible; also p_i/a_i and p_j/a_j; also a_i/a_j and p_i/p_j; also a_j/a_i and p_j/p_i." (Huntington, 1921, p. 124). Starting from this Fundamental Principle, Huntington pointed at two operationalizations —the absolute difference between such quantities is nearly zero, the ratio between them is nearly one—, and then observed that absolute difference tests are non-neutral in this respect. (recall e.g. that Dean's method and Webster's method both originate from a similar test, but with the roles of seats and votes reversed). An explicit choice of reference is avoided with relative differences and, moreover, all such relative difference tests yield the method of equal proportions.

While the thrust of the Fundamental Principle is appealing, it is not incorporated by the method of equal proportions in a voter-oriented framework. As indicated in the previous section, a concern for the roles played by seats and voters appears as the choice between either $J_\alpha(a,p)$ or $J_\alpha(p,a)$. In this respect, it is immediate from (4) that all generalized entropy measures have the property that $J_\alpha(a,p) = J_{1-\alpha}(p,a)$. It is hence also immediate that there is only one Stolarsky-mean divisor method that is neutral with respect to taking voters or seats as the reference point, and that this is *not* the method of equal proportions I_{-1}. Only when $\alpha = 1/2$ one obtains the symmetry property. Stated otherwise, only the Stolarsky-mean divisor method $I_{1/2}$ resurrects the Fundamental Principle in an internally consistent manner, so avoiding a choice between maximal equality in terms seats per voter or maximal equality in terms of electoral district sizes/voters per representative. This clearly is a non-traditional method, which to the best of our knowledge has not been discussed earlier in the literature.[12] Setting $\alpha = 1/2$ (and $\beta = 1$) in (2) clarifies that this method uses the rounding rule

$$\left(\frac{\sqrt{a+1} - \sqrt{a}}{1/2} \right)^{-2} = \left(\frac{1}{2} \left(\sqrt{a+1} + \sqrt{a} \right) \right)^2 = M_{\frac{1}{2}}(a+1, a)$$

have been emerging ever since, but they neglect that the relevant analysis pertains to absolute or relative differences between seat-to-vote *ratio's*, rather than between district sizes. In the proportional representation context, absolute measures essentially ask *how much difference* there is between seat-to-vote ratio's whereas relative differences measure *how many times* a state's ratio is bigger or smaller than that of another state. On this account, an intrinsic preference for either of the two types is difficult to state.

[12]Interestingly, the reference-distribution neutrality property has never been directly contested by the critics of the method of equal proportions, despite the fact that it was a Fundamental Principle for by Huntington. On the contrary, proponents of Webster's method have either built their case on its (statewise) unbiased character or, notably, on an explicit endorsement of the idea that voters rather than representatives *are* the natural references. Agnew (2008, p. 302) basically uses a similar 'voters-as-reference' argument to prefer J_1 (identric mean-rounding) over J_0 (logarithmic mean-rounding).

where $M_{\frac{1}{2}}(a+1, a)$ denotes the mean of order $1/2$. We recall from Lemma 1 that this divisor method is situated in between the logarithmic mean rounding and the identric mean rounding method in the favoring small states ordering and, *a fortiori*, in between the method of equal proportions and Webster's method. For completeness, we note further that there is a second sense in which it can be regarded as a specific compromise between these last two methods, since its rounding rule can be expressed as the arithmetic average of their rounding rules. That is: $\left(\frac{1}{2}\left(\sqrt{a+1} + \sqrt{a}\right)\right)^2 = \frac{1}{2}\left(\left(a + \frac{1}{2}\right) + \sqrt{a(a+1)}\right)$.

5 Concluding comments

The 'one man, one vote'-principle is an ideal of representational equality with the practical implication that voters should be given a most equal share of seats. We addressed in a general way the fundamental relation between the substantive meaning of 'most equal' and the choice among apportionment methods. We have focused indeed on the *structural* optimality properties of a class of (Stolarsky mean) divisor methods, rather than on the specific properties of one of its members. Such a general approach allows to obtain related results easily. For instance, the close connection of a subset of the generalized entropy family with Atkinson's (1970) welfare-based inequality measure can directly be exploited. In particular, Borooah's (2002) descriptive model of electoral welfare/electoral inequality can readily be extended by unveiling the links between members of the Atkinson class and divisor methods. Or, one could look at the connections between methods of apportionment and other families of objective functions, not necessarily grounded in a voter-oriented inequality assessment framework (see e.g. Balinski and Ramirez (1999), the 'error functions' surveyed by Niemeyer and Niemeyer (2008), or the multiplicative objective functions considered by Gaffke and Pükelsheim (2008)).

Such a general perspective could be criticized. As its very aim is to disclose the association between different equality objectives and their respective apportionment methods, one may dislike its ultimate relativism. Except for a brief digression at the end of section 4, we have refrained from entering a discussion that ultimately bears on conceptions of equality. It is clear that within the generalized entropy framework such discussions revolve around the choice of α (and/or the choice the reference distribution). While the inequality-aversion interpretation of α is well-established in the income inequality literature, a further clarification of this issue within the malapportionment context constitutes another interesting avenue for further research.

Conversely, one could consider the scope of our analysis too limited. Our focus on (weakly) proportional representation neglects some positive as well as normative concerns that are not without relevance. As regards the former, we have for instance ruled out existent scenario's with legal vote minimum tresholds, predefined seat maxima, or minimum representation re-

quirements.[13] Whether such constraints can be embedded in an electoral inequality analysis other than in an ad hoc fashion is an open question at this stage. The normative concerns relate to the deeper question whether seats should ideally be linearly proportional to the number of voters. It has primarily been raised in conjunction with the notion of voting power in representative democracies. Starting with Penrose (1946), this strand of literature alternatively argues that equality of ex ante voting power is achieved by allocating seats proportional to the square root of the underlying vote numbers, a recommendation that still pops up repeatedly (e.g. in recent political discussions regarding the apportionment of the enlarged European Parliament). The notion of voting power (as usually measured by the Banzhaf index) nowhere appears in our analysis. This notion requires additional information about the majority size required for decision making and about the probability distribution of individual voters' decisions. Introducing it without altering the main points of our paper is possible by assuming that voters of a specific constituency consistently act as a bloc in their voting decisions, thus replacing the random individual voting assumption upon which the square root rule is founded. This 'microcosm' interpretation of proportional representation has been sustained normatively by Thomas Paine, but seems empirically too rigid. On the other hand, the square root rule is controversial as well. Theoretically, Felsenthal and Machover (1999) pointed out that equality of voting power (in conjunction with the random voting assumption) requires the voting powers rather than the seat sizes to be proportional to the square roots of the constituency sizes. More generally, the usefulness of the a priori decision making power framework has been challenged by coalition formation models that use non-cooperative game theory (see e.g. Snyder *et al.* 2005). In addition, after empirically verifying the random voter hypothesis itself, Gelman *et al.* (2004) argue that equality of ex ante voter power often is better served by the linear proportionality ideal that we have put central. Last but not least, if one still would prefer the square root proportionality ideal, this still leaves open the practical problem of rounding ideals to integers. Ironically, that — rather than proportional representation— was precisely the problem addressed by Theil and Schrage. The framework of the current paper is, hence, readily amenable to such extensions.

A Proofs

Proof of Proposition 1. Since $D_{\alpha,\beta}$ and $D_{\beta,\alpha}$ coincide, we concentrate on the first claim. Let $\beta \neq 0$ and let $k \in \mathbb{N}$. Start from the identity

$$S_{\alpha,\beta}(k+1,k) \;=\; \left(\frac{\beta}{\alpha}\frac{(k+1)^{\alpha}-k^{\alpha}}{(k+1)^{\beta}-k^{\beta}}\right)^{1/(\alpha-\beta)} \;=\; \left(S_{\frac{\alpha}{\beta},1}((k+1)^{\beta},k^{\beta})\right)^{1/\beta}.$$

---

[13]Still, it should be recalled that all disproportionality methods I_{α} with $\alpha \leq 0$ guarantee that each state gets at least *one* seat (provided $n \leq H$).

Let $a > b$ in $\mathbb{N}$. We have to compare the fractions

$$f(\alpha', \beta; a, b) = \left(\frac{S_{\frac{\alpha'}{\beta},1}((a+1)^\beta, a^\beta)}{S_{\frac{\alpha'}{\beta},1}((b+1)^\beta, b^\beta)}\right)^{1/\beta} \quad \text{and} \quad f(\alpha, \beta; a, b) = \left(\frac{S_{\frac{\alpha}{\beta},1}((a+1)^\beta, a^\beta)}{S_{\frac{\alpha}{\beta},1}((b+1)^\beta, b^\beta)}\right)^{1/\beta}.$$

In case $\beta > 0$, then $\alpha/\beta > \alpha'/\beta$ and the map $x \mapsto x^{1/\beta}$ is strictly increasing. Lemma 1 implies that $f(\alpha'^\beta > f(\alpha, \beta; a, b)^\beta$. Hence, $f(\alpha', \beta; a, b) > f(\alpha, \beta; a, b)$. Therefore, $D_{\alpha', \beta}$ favors small states relative to $D_{\alpha, \beta}$.

In case $\beta < 0$, then $\alpha/\beta < \alpha'/\beta$ and the map $x \mapsto x^{1/\beta}$ is strictly decreasing. Lemma 1 implies that $f(\alpha'^\beta < f(\alpha, \beta; a, b)^\beta$. Hence, $f(\alpha', \beta; a, b) > f(\alpha, \beta; a, b)$. Therefore, $D_{\alpha', \beta}$ favors small states relative to $D_{\alpha, \beta}$. $\qquad\square$

Proof of Proposition 2. The method I_φ satisfies weak proportionality (equal or proportional distributions minimize Φ) and consistency (due to additive separability). In view of Balinski and Young (2001, Thm 8.4), it is sufficient to check that I_φ is weakly population monotone, that is, for each problem (p, H) and each allocation a in $I_\varphi(p, H)$, the strict inequality $p_i > p_j$ implies a weak inequality $a_i \geq a_j$. Since the disproportionality method is consistent, the couple (a_i, a_j) solves the apportionment problem $(p_i, p_j; a_i + a_j)$. The inequality $p_i > p_j$ implies $\varphi(b, p_i) < \varphi(b, p_j)$. Since the differences $\Delta\varphi(b, q)$ are increasing in b and decreasing in q, the conclusion $a_i \geq a_j$ follows. $\qquad\square$

Proof of Theorem 2. We distinguish three cases.

CASE 1. $\alpha < 0$.

We ignore the positive factor $\alpha(\alpha - 1)$ and we write $\Delta_\alpha(b, q) = [(b+1)^\alpha - b^\alpha] q^{1-\alpha}$. Plug in this expression in equilibrium condition (3) and obtain

$$\max_{i|a_i>0} \left\{ [(a_i)^\alpha - (a_i - 1)^\alpha] p_i^{1-\alpha} \right\} \leq \min_j \left\{ [(a_j + 1)^\alpha - (a_j)^\alpha] p_j^{1-\alpha} \right\}.$$

Of course, this equilibrium condition is invariant for strictly increasing transformations of the values $\Delta_\alpha(a_i, p_i)$. We bring in a minus sign (max/min becomes min/max), we divide by $-\alpha > 0$, and we apply the monotone transformation $x \mapsto x^{1/(1-\alpha)}$. We obtain, subsequently,

$$\min_{i|a_i>0} \left\{ p_i^{1-\alpha} [(a_i - 1)^\alpha - (a_i)^\alpha] \right\} \geq \max_j \left\{ p_j^{1-\alpha} [(a_i)^\alpha - (a_i + 1)^\alpha] \right\},$$

and

$$\min_{i|a_i>0} \left\{ \frac{p_i}{\left[\frac{(a_i)^\alpha - (a_i - 1)^\alpha}{\alpha} \right]^{1/(\alpha-1)}} \right\} \geq \max_j \left\{ \frac{p_j}{\left[\frac{(a_i + 1)^\alpha - (a_i)^\alpha}{\alpha} \right]^{1/(\alpha-1)}} \right\}. \tag{10}$$

The divisor sequence d_α behind I_α becomes visible: for each a in $\mathbb{N}$

$$a \leq d_\alpha(a) = \{[(a+1)^\alpha - a^\alpha]/\alpha\}^{1/(\alpha-1)} \leq a + 1.$$

We recognize the Stolarsky mean $S_{\alpha,1}(a+1,a)$ as the signpost $d_\alpha(a)$. The convention "$p_i > p_j$ implies $p_i/0 > p_j/0 > p_i$" translates towards "$p_i > p_j$ implies $-\infty \times p_i < -\infty \times p_j$" and guarantees that the larger parties receive one seat in case $H < n$.

CASE 2. $\alpha > 1$.

Start from the equilibrium condition (3) and substitute the above (case 1) expression for $\Delta_\alpha(b,q)$. Divide by α and apply the monotone decreasing transformation $x \mapsto x^{1/(1-\alpha)}$. Again, one obtains condition (10). The divisor sequence coincides with $S_{\alpha,1}$.

CASE 3. $0 < \alpha < 1$.

Now $\alpha(\alpha - 1) < 0$. We write $\iota_\alpha(b,q) = -b^\alpha q^{1-\alpha}$. Start from the equilibrium condition (3), bring in a minus sign, divide by $\alpha > 0$, and apply the monotone transformation $x \mapsto x^{1/(1-\alpha)}$. Obtain condition (10). The divisor sequence coincides with $S_{\alpha,1}$. $\qquad\square$

Proof of Corrolary 1.

CASE 1. $\alpha = 0$.

We refer to Theil and Schrage (1977). They indicate that minimizing $J_0(a,p)$ boils down to the rank-index method (or priority rule) induced by $S_{0,1}$.

CASE 2. $\alpha = 1$.

The corresponding "φ-map" is a positive affine transform of

$$\iota_1 : \mathbb{N} \times \mathbb{N}_0 \longrightarrow \mathbb{R} : (b,q) \longmapsto b\ln(b/q).$$

The difference $\Delta_1(b,q) = (b+1)\ln(b+1) - b\ln(b) - \ln(q)$ is strictly decreasing in q and strictly increasing in b. Plug the identity

$$\Delta_1(b,q) \;=\; \ln\frac{(b+1)^{b+1}}{b^b\, q},$$

in the equilibrium condition (3), bring in a minus sign, apply the monotone transformation "exp", and multiply by e. The equilibrium condition

$$\max_i \left\{ \ln\frac{a_i^{a_i}}{(a_i-1)^{a_i-1}\, p_i} \right\} \;\leq\; \min_j \left\{ \ln\frac{(a_j+1)^{a_j+1}}{a_j^{a_j}\, p_j} \right\}$$

transforms into

$$\min_i \left\{ \frac{p_i}{\frac{1}{e}\frac{a_i^{a_i}}{(a_i-1)^{a_i-1}}} \right\} \;\geq\; \max_j \left\{ \frac{p_j}{\frac{1}{e}\frac{(a_j+1)^{a_j+1}}{a_j^{a_j}}} \right\}.$$

Conclude that the divisor sequence d_1 coincides with the identric mean $S_{1,1}$. $\qquad\square$

B Inequality tests and divisor methods

We demonstrate the different nature of Huntington's unweighted state-wise inequality tests and those emerging from our framework for Webster's method, i.e. the divisor method based on arithmetic mean rounding. The min-max inequality (1) for this method takes the form

$$\frac{p_i}{a_i - \frac{1}{2}} \geq \frac{p_j}{a_j + \frac{1}{2}}. \tag{11}$$

Huntington retrieves (11) from the absolute difference test

$$a_i/p_i - a_j/p_j \gtrless (a_j + 1)/p_j - (a_i - 1)/p_i. \tag{12}$$

Clearly, in (12), whether a seat transfer would bring the states' seat-to-vote ratio's nearer to each other does not take into account how many voters would be affected. In the generalized entropy framework, a similar test for Webster's method amounts to checking whether J_2 increases if we transfer a seat between states i and j, i.e. (using (5)):

$$\frac{1}{2P} \left(p_i \left[\left(\frac{a_i/\mu_a}{p_i/\mu_p}\right)^2 - 1 \right] + p_j \left[\left(\frac{a_j/\mu_a}{p_j/\mu_p}\right)^2 - 1 \right] \right) \tag{13}$$
$$\leq \frac{1}{2P} \left(p_i \left[\left(\frac{(a_i - 1)/\mu_a}{p_i/\mu_p}\right)^2 - 1 \right] + p_j \left[\left(\frac{(a_j + 1)/\mu_a}{p_j/\mu_p}\right)^2 - 1 \right] \right).$$

It is easy to check that this test (13) eventually can be rewritten as (11). Note further that J_2, unlike (12), addresses the same voter-oriented problem that Sainte Laguë (1910) originally considered. That is, first measuring to what extent seats-per-voter differ from the overall average, where Saint Laguë proposed to measure such deviations by squaring the differences, and next weighting the individual malapportionments by the corresponding number of voters affected to arrive at an aggregate objective function.

References

[1] Agnew, R.A. (2008), Optimal Congressional Apportionment, *American Mathematical Monthly* **115,** 297-303.

[2] Balinski, M. and V. Ramirez (1999), Parametric Methods of Apportionment, Rounding and Production, *Mathematical Social Sciences* **37**, 107-122.

[3] Balinski, M.L. and H. P. Young (2001), *Fair Representation: Meeting the Ideal of One Man, One Vote* (2nd ed.), Washington D.C., Brookings Institution Press.

[4] Borooah, V. (2002), The Proportionality of Electoral Systems: Electoral Welfare and Electoral Inequality, *Economics and Politics* **14**, 83-98.

[5] Chakravarty, S. (1999), Measuring Inequality: The Axiomatic Approach, in Silber, J. (ed.), *Handbook on Inequality Measurement,* Boston, Kluwer Academic.

[6] Cowell, F. (1980), Generalized Entropy and the Measurement of Distributional Change, *European Economic Review* **13**, 147-159.

[7] Cowell, F. and K. Kuga (1981), Inequality Measurement: an Axiomatic Approach, *European Economic Review* **15**, 287-305.

[8] Felsenthal, D.S. and M. Machover (1999), Minimizing the Mean Majority Deficit: The Second Square-Root Rule, *Mathematical Social Sciences* **37**, 25-37.

[9] Gaffke, N. and F. Pukelsheim (2008), Divisor Methods for Proportional Representation Systems: An Optimization Approach to Vector and Matrix Problems, forthcoming in *Mathematical Social Sciences* (available online; doi:10.1016/j.mathsocsci.2008.01.004) .

[10] Gelman, A., Katz, J.N. and J. Bafumi (2004), Standard Voting Power Indexes Do Not Work: An Empirical Analysis, *British Journal of Political Science* **34**, 657-674.

[11] Grilli di Cortona, P., Manzi, C., Pennisi, A., Ricca, F. and B. Simeone (1999), *Evaluation and Optimization of Electoral Systems,* Philadelphia, SIAM Monographs on Discrete Mathematics.

[12] Huntington, E.V. (1921), The Mathematical Theory of the Apportionment of Representatives, *Proceedings of the National Academy of Sciences* **7**, 123-127.

[13] Huntington, E.V. (1928), The Apportionment of Representatives in Congress, *Transactions of the American Mathematical Society* **30**, 85-110.

[14] Jenkins, S. and M. O'Higgins (1989), Inequality Measurement Using 'Norm Incomes': Were Garvy and Paglin Onto Something After All?, *Review of Income and Wealth* **35**, 265-282.

[15] Marshall, A.W., Olkin I. and F. Pukelsheim (2002), A Majorization Comparison of Apportionment Methods in Proportional Representation, *Social Choice and Welfare* **19**, 885-900.

[16] Monroe, B. (1994), Disproportionality Indexes and Malapportionment: Measuring Electoral Inequity, *Electoral Studies* **13**, 132-149.

[17] Niemeyer, H.F. and A.C. Niemeyer (2008), Apportionment Methods, forthcoming in *Mathematical Social Sciences* (available online; doi:10.1016/j.mathsocsci.2008.03.003)

[18] Pennisi, A. (1998), Disproportionality Indexes and Robustness of Proportional Allocation Methods, *Electoral Studies* **17**, 3-19.

[19] Penrose, L.S. (1946), The Elementary Statistics of Majority Voting, *Journal of the Royal Statistical Society* **109**, 53-57.

[20] Sainte Laguë, M. A. (1910), La Représentation Proportionelle et la Méthode des Moindres Carrés, *Annales Scientifiques de l'École Normale Supérieure* **3**, 529-54.

[21] Schofield, N. (1982), The Relationship Between Voting and Party Strength in an Electoral System, in M. Holler (ed.), *Power, Voting, and Voting Power,* Berlin, Physica-Verlag, p. 121-134.

[22] Snyder, J.M., M.M. Ting, and S. Ansolohabere (2005), Legislative Bargaining under Weighted Voting, *American Economuc Review* **95**, 981-1004.

[23] Stolarsky, K. (1975), Generalizations of the Logarithmic Mean, *Mathematics Magazine* **48**, 87-92.

[24] Te Riele, H.J.J. (1978), The Proportional Representation Problem in the Second Chamber: an Approach via Minimal Distances, *Statistica Neerlandica* **32**, 163-179.

[25] Theil, H. (1969), The Desired Political Entropy, *American Political Science Review* **63**, 521-525.

[26] Theil, H. (1971), The Allocation of Power that Minimizes Tension, *Operations Research* **19**, 977-982.

[27] Theil, H., and Schrage, L. (1977), The Apportionment Problem and the European Parliament, *European Economic Review* **9**, 247-263.

[28] Young, H.P. (1994), *Equity in Theory and Practice*, Princeton NJ, Princeton University Press.

[29] Van Puyenbroeck, T. (forthc.), Proportional Representation, Gini Coefficients, and the Principle of Transfers, *Journal of Theoretical Politics* **20**.

[30] Willcox, W.F. (1954), Methods of Apportioning Seats in the House of Representatives, *Journal of the American Statistical Association* **49**, 685-695.