ASYMPTOTIC BIAS OF SOME ELECTION METHODS

SVANTE JANSON

Abstract. Consider an election where N seats are distributed among parties with proportions $p_1, \dots, p_m$ of the votes. We study, for the common divisor and quota methods, the asymptotic distribution, and in particular the mean, of the seat excess of a party, i.e. the difference between the number of seats given to the party and the (real) number Np_i that yields exact proportionality. Our approach is to keep $p_1, \dots, p_m$ fixed and let $N \to \infty$, with N random in a suitable way.

In particular, we give formulas showing the bias favouring large or small parties for the different election methods.

1. Introduction

The idea of a proportional election method is that each party gets a number of seats that is proportional to the number of votes. The same mathematical problem arises if seats are to be apportioned (before the election) between states in a union or between multi-member constituences according to their populations. In particular, the problem of apportionment to the House of Representatives in the United States has been the source of much debate as well as much research since 1790, including the invention of several important election methods; see Balinski and Young [1] for a detailed history. (The European Union has yet to agree on a similar procedure; see Grimmett [10] for a recent attempt to replace the current political bartering with a "mathematical formula", i.e. a well-defined method to allocate seats in the European parliament to the member states.) To fix the terminology, we talk in this paper about parties in an election; the case of apportionment between states etc. differs mathematically only in language.

Of course, exact proportionality is in general not possible, since the number of seats has to be an integer, and the election method can be seen as a method to "round" the real numbers that yield exact proportionality to integers. The purpose of this paper is to study the resulting deviation from perfect proportionality, i.e. the "rounding errors" introduced by the method. In particular, we are interested in whether there is a systematic bias favouring large or small parties for various election methods.

Date: 28 October, 2011; revised 3 April, 2012.

2000 *Mathematics Subject Classification.* 91B12; 60C05, 91F10.

Key words and phrases. proportional election methods; divisor methods; quota methods; bias; probability of violating quota; apparentements.

This problem has received a considerable amount of attention going back to (at least) Pólya [23, 24, 25]. (Sainte-Laguë [27] did a related comparison between two different methods already in 1910, see Example 3.5.) Some investigations have analysed data from real elections or from simulations (with party sizes chosen at random according to some predetermined distributions), see e.g. Balinski and Young [1] and Schuster, Pukelsheim, Drton and Draper [28]. The problem has also been studied theoretically by assuming that the parties have random sizes with a uniform distribution of the relative sizes over all possibilities, see Pólya [23, 24, 25], Schuster, Pukelsheim, Drton and Draper [28], Heinrich, Pukelsheim and Schwingenschlögl [13] and Drton and Schwingenschlögl [30, 6, 31]. Random party sizes with more general distributions have been considered by Heinrich, Pukelsheim and Schwingenschlögl [12, 13] and Schwingenschlögl [29]. (See also Section 6 below.)

The approach in the present paper is different. We consider a number of parties with fixed sizes and let the total number of seats N be random. More precisely, we let N be chosen uniformly at random in $1, 2, \dots, N_0$ for some large integer N_0, and then we take the limit as $N_0 \to \infty$. (Limits obtained in this way, if they exist, can be regarded as results for "a random positive integer" N. See also Remark 3.1.) The same approach has been used in [17] for a related problem (the Alabama paradox for Hamilton's method).

We use this approach for several different election methods, and calculate, for given sizes of the parties (under a technical condition, see Sections 3 and 11), the (asymptotic) distribution of the deviation from perfect proportionality as well as its mean, i.e. the systematic bias, and its variance. In particular, this shows how the bias depends on the sizes of the parties in a more detailed way than previous studies that consider random party sizes (see the references given above).

The election methods that we consider are divisor methods of the "linear" (or "stationary") type (including Jefferson/D'Hondt and Webster/Sainte-Laguë) and quota methods (including Hamilton/Hare and Droop). See further Section 2, and Appendices A–B.

The main results are stated in Section 3 and proven in Sections 4–5. Section 6 show that the results hold also in the more traditional approach with deterministic house size N and random party sizes. The following sections contain some simple applications of our main results: Section 7 discusses the probability of violating quota, Section 8 the expected gain or loss for parties forming an apparentement (coalition), Section 9 the Sainte-Laguë divergence studied by Heinrich, Pukelsheim and Schwingenschlögl [12, 13] and Section 10 some related goodness-of-fit functionals. Section 11 discusses our technical condition and the case of rational sizes of the parties. The appendices contain some background material.

Acknowledgement. I thank Friedrich Pukelsheim for many helpful comments.

2. Notation

We assume throughout that we have m parties with v_i votes for party i; we let $V := \sum_{i=1}^m v_i$ be the total number of votes and $p_i := v_i/V$ the proportions of votes for party i. (In real elections there may also be votes for parties that are disqualified because of a threshold rule, blank votes, and other invalid votes; these are ignored here so $\sum_{i=1}^m p_i = 1$. We consider only parties that have received at least one vote, so $p_i > 0$.) We further assume that N seats are to be distributed (the *house size*), and let s_i be the number of seats given to party i; thus $\sum_{i=1}^m s_i = N$. We occasionally write $s_i(N)$ when we need to specify the house size.

Strict proportionality would give

$$q_i := \frac{v_i}{V}N = p_i N \tag{2.1}$$

seats to party i. (This is usually not an integer.) We define the *seat excess* for party i to be the difference

$$\Delta_i := s_i - q_i = s_i - p_i N. \tag{2.2}$$

Note that

$$\sum_{i=1}^m \Delta_i = \sum_{i=1}^m s_i - N = 0. \tag{2.3}$$

We use the standard notations $\lfloor x \rfloor$ and $\lceil x \rceil$ for rounding down and up of a real number x, i.e. the largest integer $\leq x$ and the smallest integer $\geq x$, respectively. We denote the fractional part of x by $\{x\} := x - \lfloor x \rfloor$.

More generally, let α be a real number. We say that the *α-rounding* of a real number x is the integer $[x]_\alpha$ such that

$$x - \alpha \leq [x]_\alpha \leq x - \alpha + 1; \tag{2.4}$$

equivalently,

$$[x]_\alpha + \alpha - 1 \leq x \leq [x]_\alpha + \alpha. \tag{2.5}$$

If $x - \alpha$ is an integer, we regard $[x]_\alpha$ as multiple-valued, and accept both $x - \alpha$ and $x - \alpha + 1$ as values of $[x]_\alpha$. (This exceptional case typically occurs at ties in the election methods.) Consequently,

$$[x]_\alpha = \lceil x - \alpha \rceil = \lfloor x + 1 - \alpha \rfloor, \tag{2.6}$$

except the case when $x - \alpha$ is an integer and both $x - \alpha$ and $x - \alpha + 1$ are possible values. If $0 \leq \alpha \leq 1$, this means that x is rounded down if its fractional part is less than α and up if its fractional part is greater than α. In particular, $\alpha = \frac{1}{2}$ yields standard rounding, $\alpha = 0$ yields rounding up and $\alpha = 1$ yields rounding down. (But note that we allow also $\alpha < 0$ or $\alpha > 1$, in which case $|x - [x]_\alpha|$ may be greater than 1.)

The two types of election methods that we consider can be defined as follows; see further Appendices A–B where further details and equivalent definitions are given. Note that all methods that we consider are *homogeneous*, i.e., the result depends only on the proportions p_i of votes, but not on the total number V of votes; hence we will mainly use p_i rather than v_i in the sequel.

- The *β-linear divisor method*, or the *divisor method with divisors* $d(n) = n + \beta - 1$ (where β is a real number): Let
$$s_i := \left[\frac{v_i}{D}\right]_\beta = \left[\frac{p_i}{D'}\right]_\beta, \tag{2.7}$$
where D (or $D' = D/V$) is chosen such that $\sum_{i=1}^m s_i = N$. (If $\beta < 0$ or $\beta > 1$, there are some exceptions for small N, see Appendix A.1.) This includes several important methods, in particular $\beta = 1$ (Jefferson, D'Hondt), $\beta = 1/2$ (Webster, Sainte-Laguë), $\beta = 0$ (Adams).
- The *γ-quota method* (where γ is a real number): Let $Q := V/(N+\gamma)$ and let
$$s_i := \left[\frac{v_i}{Q}\right]_\alpha = [(N+\gamma)p_i]_\alpha, \tag{2.8}$$
where α is chosen such that $\sum_{i=1}^m s_i = N$. In principle, γ can be any real number (provided $N + \gamma > 0$); in practice only the cases $\gamma = 0$ (Hamilton, Hare, method of largest remainder) and $\gamma = 1$ (Droop) are important, although also $\gamma = 2$ has some limited use.

Remark 2.1. In both divisor methods and quota methods ties may occur; then the α-rounding in (2.7) or (2.8) has two possible values (because $x-\alpha \in \mathbb{Z}$ in (2.4)) for (at least) two parties, with (at least) one of them rounded to the larger value and (at least) one rounded to the smaller value. We assume that ties are resolved at random, but ties will usually not be a problem for us because of our assumptions below.

We let $\mathrm{U}(a,b)$ denote the uniform distribution on the interval (a,b). We will use U_i and $\widetilde{U}_i$ for random variables with $U_i \sim \mathrm{U}(0,1)$ and $\widetilde{U}_i \sim \mathrm{U}(-\frac{1}{2},\frac{1}{2})$, and we assume that such variables with different indices are independent. Recall that then $\mathbb{E}\, U_i = 1/2$, $\mathbb{E}\, \widetilde{U}_i = 0$ and $\operatorname{Var} U_i = \operatorname{Var} \widetilde{U}_i = 1/12$.

We denote convergence in distribution by $\xrightarrow{\mathrm{d}}$, convergence in probability by $\xrightarrow{\mathrm{p}}$ and equality in distribution by $\overset{\mathrm{d}}{=}$.

3. Main results

As explained in the introduction, we assume that $p_1, \dots, p_m$ are fixed, and let the total number of seats (house size) N be random and uniformly distributed on $\{1, \dots, N_0\}$ for some large integer N_0; we then consider limits as $N_0 \to \infty$. For convenience, we denote this situation by $N \xrightarrow{\mathrm{p}*} \infty$. (Note that this implies $N \xrightarrow{\mathrm{p}} \infty$.)

Remark 3.1. We assume that N is distributed in this way for simplicity, but the results can easily be extended to more general sequence of random $N \in \mathbb{N}$. For example, we may take N uniformly distributed on an interval $[N_1(k), N_2(k)]$ and let $k \to \infty$, for any two sequences $N_1(k)$ and $N_2(k)$ of positive integers with $N_2(k) - N_1(k) \to \infty$. In fact, we conjecture that the results hold for any N that satisfies the weak condition

$$\varphi(\theta) := \mathbb{E}\, e^{\mathrm{i}N\theta} \to 0 \qquad \text{for any } \theta \in (0, 2\pi). \tag{3.1}$$

(This implies $N \xrightarrow{\mathrm{p}} \infty$ and excludes parity restrictions.) Note that Weyl's Lemma 4.1 extends to this situation. However, we have not verified all details for this case, and we leave further investigations to the interested reader.

In order to obtain simple asymptotic results, we will usually make one mathematical simplification: we assume that $p_1, \dots, p_m$ are *linearly independent over* $\mathbb{Q}$, i.e., that there is no relation

$$a_1 p_1 + \dots + a_m p_m = 0 \tag{3.2}$$

with all coefficients a_i rational and not all 0. (Equivalently, there is no relation (3.2) with integer coefficients a_i, not all 0.) The case when this assumption does not hold is discussed in Section 11.

Remark 3.2. Mathematically, this assumption is reasonable, since if we choose $p_1, \dots, p_m$ at random (uniformly given that their sum is 1, as in Section 6), they will almost surely be linearly independent over $\mathbb{Q}$. However, for real elections the assumption is clearly unreasonable since vote counts v_i are integers and the p_i thus rational numbers. Nevertheless, the results below are good approximations if the numbers p_i have large denominators and there are no relations (3.2) with small integers $a_1, \dots, a_m$, see Section 11. For practical purposes we thus can use the results below as good approximations for any p_i and any N that is large enough. (However, we do not investigate the rate of convergence, and how large N has to be.)

3.1. **Divisor methods.** We begin with a simple deterministic bound, valid for all house sizes and numbers of votes, cf. [18, Satz 6.2.11]. It follows from the proof, or from Theorem 3.7 below, that the bounds are best possible. (Among bounds not depending on N; for a given N, the lower bound can be improved when N is so small that not all parties get seats, see Kopfermann [18, p. 131].) Proofs of the results below are given in Section 4.

Theorem 3.3. *Consider the β-linear divisor method. Then*

$$p_i - 1 + (\beta - 1)\,(mp_i - 1) \le \Delta_i \le (\beta - 1)\,(mp_i - 1) + (m-1)p_i. \tag{3.3}$$

Equivalently,

$$\left|\Delta_i - \Bigl(\beta - \frac{1}{2}\Bigr)(mp_i - 1)\right| \le \frac{1}{2} + \frac{m-2}{2} p_i. \tag{3.4}$$

For the asymptotic bias, we have a simple formula.

Theorem 3.4. *Consider the β-linear divisor method, and suppose that p_1, $\dots, p_m$ are linearly independent over $\mathbb{Q}$. Then, as $N \xrightarrow{\mathrm{p}*} \infty$,*

$$\mathbb{E}\,\Delta_i \to \bigl(\beta - \tfrac{1}{2}\bigr)\,(mp_i - 1). \tag{3.5}$$

Note that the asymptotic bias in (3.5) for a party depends only on its size p_i and the number of parties, but not on the distribution of sizes between the other parties. The limit in (3.5) agrees with the formula by Kopfermann [18, (S) on p. 208] for the advantage (or disadvantage) for different parties, derived by geometric considerations.

In particular, Theorem 3.4 shows that the bias is 0 for every party when $\beta = 1/2$ (Webster/Sainte-Laguë). It is well-known that this divisor method is (asymptotically) unbiased, see e.g. [25, 1, 28, 30, 6] for various justifications of this; it is satisfying that our approach confirms this, and shows that the method really is unbiased for a party of any size.

For $\beta > 1/2$ (for example Jefferson/D'Hondt with $\beta = 1$), the bias is positive for $p_i > 1/m$ and negative for $p_i < 1/m$. It is well-known that these methods favour larger parties; we here see exactly how much parties of different sizes are favoured, and that the boundary is exactly at the average size $1/m$. Note also that the bias grows with m, especially for large parties. For example, for $\beta = 1$, a large party with $p_i \approx 1/2$ has bias $\approx (m-2)/4$, while a small party ($p_i \ll 1/m$) always has a bias $\approx -1/2$.

For $\beta < 1/2$ (for example Adams with $\beta = 0$), we have the opposite biases; the boundary is still at $p_i = 1/m$.

Example 3.5. Sainte-Laguë [27] compared in 1910 the results of his method (which he called the *method of least squares*) and D'Hondt's method in the case of two parties (that he thought of as two of several parties in an election) with $p_1 \geq p_2$. He did not consider the biases in our sense, comparing the distribution of the seats to the distribution of the votes, but his approach means studying the difference of the two biases. (Since we know that Sainte-Laguë's method ($\beta = 1/2$) is unbiased, this is equivalent to calculating the bias of D'Hondt's method ($\beta = 1$).)

Sainte-Laguë found that if $k := p_2/p_1 \in (0, 1]$, then the largest party gets on the average $(1-k)/(2(1+k))$ seats more with D'Hondt's method. This agrees with Theorem 3.4, with $m = 2$ and $p_1 = 1/(1+k)$.

Sainte-Laguë [27] also considered the case of random p_i, assuming that the ratio p_2/p_1 is uniformly distributed over $(0, 1]$, and found by integration the average gain

$$\int_0^1 \frac{1-k}{2(1+k)}\,\mathrm{d}k = \log 2 - \frac{1}{2} \approx 0.193 \tag{3.6}$$

for the larger party with D'Hondt's method. (Note that this differs from the average bias $1/4 = 0.25$ for the largest of two parties derived by [28] assuming that p_2 is uniformly distributed on $(0, 1/2)$, see Theorem 6.4 below with $m = 2$. The reason is that the two probability distributions of the party sizes are different; they are uniform in different ways.)

Remark 3.6. Sainte-Laguë [27] seems to have thought that this calculation for two parties applies to any two of several parties, and that therefore, if party A is larger than B, then B would on the average gain 0.193 seats compared to A if the method is changed from D'Hondt's to Sainte-Laguë's, regardless of the number and sizes of the other parties. Theorem 3.4 shows that this is incorrect; in fact, if A and B both are small parties, they both have essentially the same bias $-1/2$ with D'Hondt's method. It is true that for divisor methods, the distribution of seats to two parties A and B is the same as if the total number of seats for these two parties is distributed between them by the same divisor method, see Theorem A.19. However, if both A and B are small parties, they will by (3.5) get on the average almost 1 seat less together than their proportional share, and if their total number is redistributed between them, most of this loss will be borne by the larger party, which offsets the advantage of the larger party in the redistribution and explains the fallacy. More generally, parties 1 and 2 will by (3.5) on the average get together $N' := N(p_1+p_2)+(\beta-1/2)(mp_1+mp_2-2)$ seats. If we distribute this number of seats between the two parties, the same formula would yield that party 1 on the average gets, with $p_1' := p_1/(p_1+p_2)$,

$$\begin{aligned} N'p_1' + (\beta-\tfrac12)(2p_1'-1) &= Np_1 + (\beta-\tfrac12)\bigl(p_1'(m(p_1+p_2)-2)+2p_1'-1\bigr) \\ &= Np_1 + (\beta-\tfrac12)(mp_1-1) \end{aligned}$$

seats, in accordance with Theorem 3.4.

We can describe asymptotically not only the mean but also the distribution of the seat excess. We also obtain a joint limit distribution, which is explicit although a little involved.

Theorem 3.7. *Consider the β-linear divisor method, and suppose that $p_1,\dots,p_m$ are linearly independent over $\mathbb{Q}$. Then, as $N \xrightarrow{\text{p}*} \infty$, for each i,*

$$\Delta_i \xrightarrow{\text{d}} \bar{X}_i := (\beta-\tfrac12)(mp_i-1) + \widetilde{U}_0 + p_i\sum_{k=1}^{m-2}\widetilde{U}_k, \tag{3.7}$$

where $\widetilde{U}_k \sim \mathrm{U}(-\frac12,\frac12)$ are independent. Moreover, jointly for $i=1,\dots,m$,

$$\Delta_i \xrightarrow{\text{d}} X_i, \tag{3.8}$$

where the limit random variables X_i can be constructed as follows: Let $U_1,\dots,U_m \sim \mathrm{U}(0,1)$ and let $J \in \{1,\dots,m\}$ be random with $\mathbb{P}(J=j)=p_j$, $j=1,\dots,m$, with all variables independent. Let

$$V_i := \mathbf{1}\{J\neq i\}U_i = \begin{cases} U_i, & i \neq J, \\ 0, & i = J, \end{cases} \tag{3.9}$$

and let, finally,

$$X_i := p_i\sum_{j=1}^{m} V_j - V_i + (\beta-1)(mp_i-1). \tag{3.10}$$

Of course, as we shall verify in Section 4, $X_i \overset{\text{d}}{=} \bar{X}_i$.

The description in (3.9)–(3.10) of the joint distribution of $X_1, \dots, X_m$ is explicit but rather involved, and it is quite possible that there exists a simpler description; cf. Remark 10.3. For $m = 2$, $X_2 = -X_1$, and thus $(X_1, X_2) \overset{\mathrm{d}}{=} (\bar{X}_1, -\bar{X}_1)$ where $\bar{X}_1 := (2\beta - 1)p_1 - \beta + U_0$ by (3.7). We leave it as an open problem to try to find alternative expressions for $m \geq 3$.

Corollary 3.8. *Consider the β-linear divisor method, and suppose that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. Then, as $N \xrightarrow{\mathrm{p}*} \infty$,*

$$\operatorname{Var} \Delta_i \to \frac{1 + (m-2)p_i^2}{12} \tag{3.11}$$

and, when $i \neq j$,

$$\operatorname{Cov}(\Delta_i, \Delta_j) \to \frac{(m-2)p_i p_j - p_i - p_j}{12}. \tag{3.12}$$

Note that the asymptotic distribution of the seat excess depends on β only through its mean $\mathbb{E}\, X_i$ (i.e., the asymptotic bias in (3.5)); the centred random variable $X_i - \mathbb{E}\, X_i$ does not depend on β. We see also, from (3.10) or (3.11), that this random part of the seat excess tends to be larger in absolute value for a large party.

In the limit as $p_i \to 0$, i.e. for very small parties, (3.9) and (3.10) yield $V_i \xrightarrow{\mathrm{d}} U_i$ and

$$X_i \xrightarrow{\mathrm{d}} -U_i - (\beta - 1) = 1 - U_i - \beta \overset{\mathrm{d}}{=} U_i - \beta \sim \mathrm{U}(-\beta, 1 - \beta).$$

Remark 3.9. A simple heuristic motivation for the bias in Theorem 3.4 is that the β-rounding in (2.7) is unbiased if $\beta = 1/2$, but otherwise, it costs on the average each party $\beta - 1/2$ seats, compared to v_i/D which is truly proportional. In total, this makes a loss of $m(\beta - 1/2)$ seats, which are redistributed proportionally (by shifting D), so party i gets back $p_i m(\beta - 1/2)$ seats, on the average, giving a net gain of $(mp_i - 1)(\beta - 1/2)$, as shown by Theorem 3.4.

For a related, but rigorous, argument, let us compare the methods with parameters β and $\beta + 1$, temporarily using a superscript to distinguish the methods. (For example, Adams's method and Jefferson's method.) By Theorem A.12 (assuming that N is large enough),

$$s_i^{(\beta)}(N) = s_i^{(\beta+1)}(N - m) + 1. \tag{3.13}$$

Hence,

$$\Delta_i^{(\beta)}(N) = s_i^{(\beta)}(N) - Np_i = \Delta_i^{(\beta+1)}(N - m) - mp_i + 1. \tag{3.14}$$

Taking the limit as $N \xrightarrow{\mathrm{p}*} \infty$, we see that the asymptotic distributions also differ by $mp_i - 1$. Hence, it is not surprising that the bias in Theorem 3.4 is $mp_i - 1$ times a linear function of β.

3.2. **Quota methods.** We again begin with a simple deterministic bound, cf. [18, Satz 6.2.3]. It follows from the proof, or from Theorem 3.13 below, that the bounds are best possible. Proofs of the results below are given in Section 5.

Theorem 3.10. *For the γ-quota method,*

$$\gamma\Bigl(p_i - \frac{1}{m}\Bigr) - \frac{m-1}{m} \le \Delta_i \le \gamma\Bigl(p_i - \frac{1}{m}\Bigr) + \frac{m-1}{m}. \tag{3.15}$$

Again we have a simple formula for the asymptotic bias.

Theorem 3.11. *Consider the γ-quota method and suppose that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. Then, as $N \xrightarrow{\text{p*}} \infty$,*

$$\mathbb{E}\,\Delta_i \to \gamma\Bigl(p_i - \frac{1}{m}\Bigr). \tag{3.16}$$

As for the divisor methods, the asymptotic bias in (3.16) for a party depends only on its size p_i and the number of parties, but not on the distribution of sizes between the other parties. The limit in (3.16) agrees with the formula by Kopfermann [18, (S) on p. 199] for the (dis)advantage for different parties, derived by geometric considerations.

In particular, Theorem 3.11 shows that the bias is 0 for every party when $\gamma = 0$ (method of largest remainder/Hamilton/Hare). Again it is well-known that this quota method is (asymptotically) unbiased, see e.g. [25, 1, 28, 30, 6]; our approach confirms this and shows that the method is unbiased for a party of any size.

For $\gamma = 1$ (Droop), the bias is positive for $p_i > 1/m$ and negative for $p_i < 1/m$, just as for Jefferson's method discussed above; hence Droop's method too favours larger parties. (See [18, p. 118] for figures illustrating the cases $m = 2$ and $m = 3$.) Note, however, that the bias for Droop's method does not grow with m. A comparison between (3.16) (with $\gamma = 1$) and (3.5) (with $\beta = 1$) shows that the bias for Jefferson's method is $m/2$ times the bias for Droop's method, for any party of any size. Hence Droop's method is less biased than Jefferson's for every $m \ge 3$. (For $m = 2$ not only the biases are the same; in fact the methods coincide when there are only two parties, see Appendix B.1.)

Remark 3.12. There is a simple heuristic explanation of the formula (3.16) for the bias, if we accept the fact that Hamilton/Hare's method ($\gamma = 0$) is unbiased. Consider $\gamma = 1$ (Droop); this method can be seen as doing Hamilton's method with $N + 1$ seats, but retracting the last seat. Since Hamilton's method is unbiased, this first distributes on the average $(N+1)p_i$ seats to party i. The retracted seat is taken essentially uniformly at random, since it depends on the fractional parts of the numbers $q_i = Np_i$ only; hence each party loses on the average $1/m$ seats in that step. This argument gives a bias of $p_i - 1/m$, in accordance with (3.16).

Again we have an explicit expression also for the asymptotic distribution of the seat excess.

Theorem 3.13. *Consider the γ-quota method and suppose that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. Then, as $N \xrightarrow{\text{p*}} \infty$, for each i,*

$$\Delta_i \xrightarrow{\text{d}} Y_i := \gamma\Big(p_i - \frac{1}{m}\Big) + \widetilde{U}_0 + \frac{1}{m}\sum_{k=1}^{m-2} \widetilde{U}_k, \tag{3.17}$$

where $\widetilde{U}_k \sim \mathrm{U}(-\frac{1}{2}, \frac{1}{2})$ are independent.

Corollary 3.14. *Consider the γ-quota method and suppose that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. Then, as $N \xrightarrow{\text{p*}} \infty$,*

$$\operatorname{Var} \Delta_i \to \frac{1 + (m-2)/m^2}{12} = \frac{(m+2)(m-1)}{12m^2}. \tag{3.18}$$

As for the divisor methods in Theorem 3.7, the asymptotic distribution of the seat excess depends on γ only through the mean in (3.16); $Y_i - \mathbb{E}\, Y_i$ does not depend on γ. Moreover, unlike the case of divisor methods, this centred random variable does not depend on the party size p_i. In particular, for Hamilton/Hare's method ($\gamma = 0$), the asymptotic distribution of the seat excess is the same for every party.

Remark 3.15. Let for simplicity $\gamma = 0$. Note that the limit variable Y_i in (3.17) is uniform $U(-\frac{1}{2}, \frac{1}{2})$ for $m = 2$; for $m \geq 3$ the distribution is more complicated, but as $m \to \infty$, the limit converges to $U(-\frac{1}{2}, \frac{1}{2})$ (by (3.17) and the law of large numbers). Hence, for large m, Y_i has essentially the same distribution as for $m = 2$, and although the supremum and infimum of the range of Y_i tend to 1 and -1 as $m \to \infty$, cf. (3.15), the probability $\mathbb{P}(Y_i \notin [-\frac{1}{2}, \frac{1}{2}]) \to 0$.

Similarly, it is seen from the variance formula (3.18) that the variance is $1/12 \approx 0.08333$ for $m = 2$, increases to $5/54 \approx 0.09259$ for $m = 3$ and to the maximum $3/32 = 0.09375$ for $m = 4$; the variance then decreases again and converges to $1/12$ as $m \to \infty$.

The difference from a uniform variable is not very large, and it may in practice be a good approximation to regard the seat excess Δ_i in Hamilton's method as $\mathrm{U}(-\frac{1}{2}, \frac{1}{2})$ for any party and any reasonably large house size.

We have stated Theorem 3.13 only for individual parties. It is easy to see that a joint asymptotic distribution exists, but (unlike the case of divisor methods, cf. Theorem 3.7) we do not know any simple description of it. (There is a description implicit in the proof of Theorem 3.16, see (5.14).)

Theorem 3.16. *Consider the γ-quota method and suppose that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. Then, as $N \xrightarrow{\text{p*}} \infty$, $\Delta_i \xrightarrow{\text{d}} \bar{Y}_i$ jointly for $i = 1, \dots, m$ for some random variables $\bar{Y}_i$ such that the distribution of $(\bar{Y}_1 - \gamma p_1, \dots, \bar{Y}_m - \gamma p_m)$ is symmetric (exchangeable) and independent of $p_1, \dots, p_m$, and $\bar{Y}_i \overset{\text{d}}{=} Y_i$ for each i.*

Corollary 3.17. *Consider the γ-quota method and suppose that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. Then, as $N \xrightarrow{\mathrm{p*}} \infty$, when $i \neq j$,*

$$\operatorname{Cov}(\Delta_i, \Delta_j) \to \operatorname{Cov}(\bar{Y}_i, \bar{Y}_j) = -\frac{\operatorname{Var} \bar{Y}_i}{m-1} = -\frac{m+2}{12m^2}. \tag{3.19}$$

Problem 3.18. Find an explicit formula for the limit variables $\bar{Y}_1, \dots, \bar{Y}_m$ in Theorem 3.16.

4. Proofs for divisor methods

Proof of Theorems 3.3 and 3.7. For the β-linear divisor method, the seat distribution is given by (2.7) with D and $D' = D/V$ such that $\sum_{i=1}^m s_i = N$. We let $D' = 1/t$; thus, using (2.7) and (2.6),

$$s_i = [p_i t]_\beta = \lfloor p_i t - \beta + 1 \rfloor \tag{4.1}$$

(with the usual possible exception if $p_i t - \beta$ is an integer). Let here t grow from 0 to ∞; think of t as time. The seats are assigned one by one (unless there is a tie).

Fix one party j and let t_{jr} be the time when party j receives its r:th seat. By (4.1), $r = p_j t_{jr} - \beta + 1$, and thus

$$t_{jr} = \frac{r + \beta - 1}{p_j}. \tag{4.2}$$

At this time t_{jr}, by (4.1) again, party i has s_i seats with

$$s_i = \lfloor p_i t_{jr} - \beta + 1 \rfloor = \left\lfloor \frac{p_i}{p_j}(r + \beta - 1) - \beta + 1 \right\rfloor = \left\lfloor \frac{p_i}{p_j} r + \Bigl(\frac{p_i}{p_j} - 1\Bigr)(\beta - 1) \right\rfloor. \tag{4.3}$$

Let, for $i = 1, \dots, m$,

$$W_i := \{p_i t_{jr} - \beta + 1\} = \left\{ \frac{p_i}{p_j} r + \Bigl(\frac{p_i}{p_j} - 1\Bigr)(\beta - 1) \right\} \in [0, 1], \tag{4.4}$$

defining $W_i := 1$ in the exceptional case when $p_i t_{jr} - \beta \in \mathbb{Z}$ and $s_i = p_i t_{jr} - \beta$. (But $W_i := 0$ when $p_i t_{jr} - \beta \in \mathbb{Z}$ and $s_i = p_i t_{jr} - \beta + 1$.) Then (4.3) can be written

$$s_i = \frac{p_i}{p_j} r + \Bigl(\frac{p_i}{p_j} - 1\Bigr)(\beta - 1) - W_i, \qquad i = 1, \dots, m. \tag{4.5}$$

Note that $W_j = 0$, since $s_j = r = p_j t_{jr} - \beta + 1$.

The total number of seats at the moment t_{jr} is, by (4.5),

$$N = \sum_{i=1}^m s_i = \frac{1}{p_j} r + \Bigl(\frac{1}{p_j} - m\Bigr)(\beta - 1) - \sum_{i=1}^m W_i, \tag{4.6}$$

and thus the seat excess is, combining (4.5) and (4.6),

$$\Delta_i = s_i - Np_i = \Bigl(\frac{p_i}{p_j} - 1\Bigr)(\beta - 1) - W_i - \Bigl(\frac{p_i}{p_j} - mp_i\Bigr)(\beta - 1) + p_i \sum_{k=1}^{m} W_k$$
$$= p_i \sum_{k=1}^{m} W_k - W_i + (\beta - 1)(mp_i - 1). \tag{4.7}$$

So far everything is deterministic. Theorem 3.3 follows from (4.7) since each $W_k \in [0,1]$ and $W_j = 0$.

Let us now assume that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$; in particular, this implies that p_i/p_j is irrational when $i \neq j$. Note that there is a tie between parties i and j at time t if and only if both $p_i t - \beta$ and $p_j t - \beta$ are integers. If parties i and j tie at two different times t_1 and t_2, we thus have $p_i(t_1 - t_2) \in \mathbb{Z}$ and $p_j(t_1 - t_2) \in \mathbb{Z}$, and taking the quotient we find $p_i/p_j \in \mathbb{Q}$, a contradiction. Hence there is at most one tie for each pair of parties, and thus at most $\binom{m}{2}$ values of t (and thus of N) for which there is any tie. Since we consider asymptotics, we may ignore the ties and assume that no tie occurs.

Let now N be random in $\{1, \dots, N_0\}$ and let $J = J(N)$ be the party getting the N:th seat. Then Δ_i is given by (4.7), where $W_i = W_i(N)$ for each N is given by (4.4) with $j = J$ and $r = s_J(N)$. Fix again j and consider first only N such that $J = j$. The number of such $N \leq N_0$ is $s_j(N_0)$, the number of seats party j receives when the house size is N_0; note that, e.g. by Theorem 3.3, $s_j(N_0) = p_j N_0 + O(1)$.

We recall a well-known result by Weyl [32]. (The result is usually stated with $a_1 = \dots = a_k = 0$; this case implies immediately the more general version given here.) We say that an infinite sequence $(v_n)_{n\geq 1} \in [0,1)^k$ is *uniformly distributed* if the empirical distributions $n_0^{-1} \sum_{n=1}^{n_0} \delta_{v_n}$ converge to the uniform distribution as $n_0 \to \infty$, where δ_{v_n} denotes the Dirac measure. (Recall that this means that if $A \subseteq [0,1)^k$ with $\lambda(\partial A) = 0$, then $\#\{n \leq n_0 : v_n \in A\}/n_0 \to \lambda(A)$, where λ is the usual Lebesgue measure, see e.g. [2].)

Lemma 4.1 (Weyl)**.** *Suppose that $y_1, \dots, y_k$ and 1 are linearly independent over $\mathbb{Q}$, and let $a_1, \dots, a_k$ be any real numbers. Then the sequence of vectors $(\{ny_1 + a_1\}, \dots, \{ny_k + a_k\})_{n\geq 1} \in [0,1)^k$ is uniformly distributed in $[0,1)^k$.* □

(The standard proof is by showing that the Fourier transform (characteristic function) $\frac{1}{n_0} \sum_{n=1}^{n_0} \exp\bigl(2\pi \mathrm{i} \sum_{j=1}^{k} \ell_j \{ny_j\}\bigr) \to 0$, as $n_0 \to \infty$, for any fixed integers $\ell_1, \dots, \ell_k$, not all 0, see for example [9, Exercises 3.4.2–3].)

Since $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$, so are $p_1/p_j, \dots, p_m/p_j$, i.e. $\{p_i/p_j : i \neq j\} \cup \{1\}$. Taking $y_i := p_i/p_j$ and $a_i := (y_i - 1)(\beta - 1)$, we thus see from Lemma 4.1 that $(W_i)_{i\neq j}$, given by (4.4) for $r = 1, \dots, s_j(N_0)$, are uniformly distributed vectors in $[0,1)^{m-1}$. In probabilistic notation, regarding $W_i = W_i(N)$ as random variables, conditioning on $J = j$ and recalling $W_j = 0$ when $J = j$, as $N_0 \to \infty$ and thus $s_j(N_0) \to \infty$,

$$\big((W_1,\dots,W_m) \mid J = j\big) \xrightarrow{\mathrm{d}} \big(\mathbf{1}\{i \neq j\} U_i\big)_{i=1}^m. \tag{4.8}$$

Since $\mathbb{P}(J = j) = s_j(N_0)/N_0 \to p_j$ as $N_0 \to \infty$, the random variable J has asymptotically the distribution given in Theorem 3.7, and it follows from (4.8) that, with V_i defined as in Theorem 3.7,

$$(W_1,\dots,W_m) \xrightarrow{\mathrm{d}} (V_1,\dots,V_m). \tag{4.9}$$

Consequently, (4.7) yields $\Delta_i \xrightarrow{\mathrm{d}} X_i$, jointly for all $i = 1,\dots,m$, with X_i defined by (3.10).

To obtain the simpler form (3.7) for an individual Δ_i, let

$$\widetilde{X}_i := X_i - (\beta - \tfrac12)(mp_i - 1) = p_i \sum_{j=1}^m V_j - V_i - \tfrac12 mp_i + \tfrac12. \tag{4.10}$$

Then, if $J = i$,

$$\widetilde{X}_i = p_i \sum_{j\neq i} U_j - \tfrac12 mp_i + \tfrac12 \overset{\mathrm{d}}{=} p_i \sum_{j=1}^{m-1} \widetilde{U}_j + \tfrac12 - \tfrac12 p_i,$$

and if $J \neq i$,

$$\widetilde{X}_i \overset{\mathrm{d}}{=} p_i \sum_{j\neq i,J} U_j + (p_i - 1)U_i - \tfrac12 mp_i + \tfrac12 \overset{\mathrm{d}}{=} p_i \sum_{j=1}^{m-2} \widetilde{U}_j + (p_i - 1)\widetilde{U}_0 - \tfrac12 p_i.$$

Hence, if we define

$$Z := \begin{cases} p_i \widetilde{U}_{m-1} + \tfrac12 - \tfrac12 p_i, & J = i, \\ (p_i - 1)\widetilde{U}_0 - \tfrac12 p_i, & J \neq i, \end{cases} \tag{4.11}$$

then

$$\widetilde{X}_i \overset{\mathrm{d}}{=} p_i \sum_{j=1}^{m-2} \widetilde{U}_j + Z,$$

with $\widetilde{U}_j$ $(1 \le j \le m-2)$ and Z independent. By (4.11), conditioning on J, we have the distributions $(Z \mid J = i) \sim \mathrm{U}(\frac12 - p_i, \frac12)$ and $(Z \mid J \neq i) \sim \mathrm{U}(-\frac12, \frac12 - p_i)$; furthermore $\mathbb{P}(J = i) = p_i$, and it follows that $Z \sim \mathrm{U}(-\frac12, \frac12)$, so $Z \overset{\mathrm{d}}{=} \widetilde{U}_0$ and thus

$$\widetilde{X}_i \overset{\mathrm{d}}{=} p_i \sum_{j=1}^{m-2} \widetilde{U}_j + \widetilde{U}_0.$$

The definition (4.10) now shows that $X_i \overset{\mathrm{d}}{=} \bar{X}_i$, with $\bar{X}_i$ defined in (3.7). □

Proof of Theorem 3.4 and Corollary 3.8. Since Δ_i is bounded, e.g. by Theorem 3.3 (or by the simpler Theorem A.17), the convergence in distribution $\Delta_i \xrightarrow{\mathrm{d}} X_i$ in Theorem 3.7 implies convergence of all moments and mixed moments, see [11, Theorem 5.5.9]; hence it suffices to calculate the expectations, variances and covariances of X_i.

We have $\mathbb{E}\, U_i = 1/2$ and thus by (3.9), since J and U_i are independent,

$$\mathbb{E}\, V_i = \mathbb{P}(J \neq i)\, \mathbb{E}\, U_i = (1 - p_i)\tfrac{1}{2}. \tag{4.12}$$

Consequently, $\mathbb{E} \sum_{i=1}^m V_i = \frac{1}{2} \sum_{i=1}^m (1 - p_i) = \frac{1}{2}(m-1)$ and, by (3.10),

$$\begin{aligned}
\mathbb{E}\, X_i &= p_i\, \mathbb{E} \sum_{j=1}^m V_j - \mathbb{E}\, V_i + (\beta - 1)(mp_i - 1) \\
&= p_i \frac{m-1}{2} - \frac{1-p_i}{2} + (\beta - 1)(mp_i - 1) = \Bigl(\beta - \frac{1}{2}\Bigr)(mp_i - 1).
\end{aligned}$$

This shows Theorem 3.4.

For the variance and covariances, we have for any i

$$\mathbb{E}\, V_i^2 = \mathbb{P}(J \neq i)\, \mathbb{E}\, U_i^2 = (1 - p_i)\tfrac{1}{3}$$

and, when $i \neq j$, by independence,

$$\mathbb{E}(V_i V_j) = \mathbb{E}\bigl(\mathbf{1}\{J \neq i\}\mathbf{1}\{J \neq j\} U_i U_j\bigr) = \mathbb{P}(J \notin \{i,j\})\, \mathbb{E}\, U_i\, \mathbb{E}\, U_j = \frac{1 - p_i - p_j}{4}.$$

Consequently, using also (4.12),

$$\begin{aligned}
\operatorname{Var}(V_i) &= \mathbb{E}\, V_i^2 - (\mathbb{E}\, V_i)^2 = \frac{1-p_i}{3} - \frac{(1-p_i)^2}{4} = \frac{1 + 2p_i - 3p_i^2}{12}, \\
\operatorname{Cov}(V_i, V_j) &= \frac{1 - p_i - p_j}{4} - \frac{1-p_i}{2} \cdot \frac{1-p_j}{2} = -\frac{p_i p_j}{4} = -\frac{3p_i p_j}{12}.
\end{aligned}$$

We obtain from this also

$$\begin{aligned}
\operatorname{Cov}\Bigl(V_i, \sum_{j=1}^m V_j\Bigr) &= \sum_{j=1}^m \operatorname{Cov}(V_i, V_j) = \frac{1+2p_i}{12} - \sum_{j=1}^m \frac{3p_i p_j}{12} \\
&= \frac{1+2p_i}{12} - \frac{3p_i}{12} = \frac{1-p_i}{12}, \\
\operatorname{Var}\Bigl(\sum_{i=1}^m V_i\Bigr) &= \sum_{i=1}^m \operatorname{Cov}\Bigl(V_i, \sum_{j=1}^m V_j\Bigr) = \sum_{i=1}^m \frac{1-p_i}{12} = \frac{m-1}{12}.
\end{aligned}$$

Consequently, (3.10) yields

$$\begin{aligned}
\operatorname{Var}(X_i) &= p_i^2 \operatorname{Var}\Bigl(\sum_{j=1}^m V_j\Bigr) - 2p_i \operatorname{Cov}\Bigl(V_i, \sum_{j=1}^m V_j\Bigr) + \operatorname{Var}(V_i) \\
&= p_i^2 \frac{m-1}{12} - 2p_i \frac{1-p_i}{12} + \frac{1 + 2p_i - 3p_i^2}{12} = \frac{(m-2)p_i^2 + 1}{12}
\end{aligned}$$

and, for $i \neq j$,

$$\begin{aligned}
\operatorname{Cov}(X_i, X_j) = p_i p_j \operatorname{Var}\Bigl(\sum_{k=1}^m V_k\Bigr) - p_i \operatorname{Cov}\Bigl(V_j, \sum_{k=1}^m V_k\Bigr) - p_j \operatorname{Cov}\Bigl(V_i, \sum_{k=1}^m V_k\Bigr) \\
+ \operatorname{Cov}(V_i, V_j)
\end{aligned}$$

$$= p_i p_j \frac{m-1}{12} - \frac{p_i(1-p_j)+p_j(1-p_i)}{12} - \frac{3p_ip_j}{12}$$
$$= \frac{(m-2)p_ip_j - p_i - p_j}{12}.$$

Corollary 3.8 follows. □

Remark 4.2. We can also see when equality holds in Theorem 3.3. By (4.7), Δ_i is maximal when $W_i = 0$ and $W_k = 1$ for $k \neq i$. Thus all $W_k \in \{0, 1\}$, which by (4.4) means that all $p_k t - \beta$ are integers for $t = t_{jr}$; hence all parties tie in the sense that there are two possible values of s_i in the β-rounding in (4.1), or equivalently (2.7), for every party. Consequently Δ_i equals the upper bound in (3.3) if and only if all parties tie and, moreover, party i is lucky and is rounded up while all others are rounded down. (In the sequential formulation D3 or D4 in Appendix A, this means that all parties tie for the last seat, which is taken by party i.)

Similarly, Δ_i is minimal when $W_i = 1$ and $W_k = 0$ for $k \neq i$, and it follows that Δ_i equals the lower bound in (3.3) if and only if all parties tie and, moreover, party i is unlucky and is rounded down while all others parties are rounded up. (In the sequential formulation this means that all parties tie for the last $m-1$ seats, which are given to all parties except i.)

5. Proofs for quota methods

Proof of Theorems 3.10 and 3.13. The proof for quota methods uses arguments similar to the proof in [17], and the reader might wish to compare the versions. (Only the case $\gamma = 0$ is treated in [17], but that is only a minor simplification. For completeness we repeat some results from [17].) For notational convenience, we show the results for party 1, i.e. we take $i = 1$ in the proofs.

By definition, $s_j = [(N+\gamma)p_j]_\alpha$ for an α that makes $\sum_{j=1}^m s_j = N$. Explicitly, by (2.5),

$$s_j + \alpha - 1 \le (N+\gamma)p_j \le s_j + \alpha. \tag{5.1}$$

Thus, using party 1 as a benchmark,

$$s_j - s_1 - 1 \le (N+\gamma)(p_j - p_1) \le s_j - s_1 + 1. \tag{5.2}$$

Since $\Delta_j - \Delta_1 = s_j - s_1 - N(p_j - p_1)$, (5.2) yields

$$\Delta_j - \Delta_1 - 1 \le \gamma(p_j - p_1) \le \Delta_j - \Delta_1 + 1. \tag{5.3}$$

Thus, summing over all $j \neq 1$, recalling that $\sum_{j=1}^m \Delta_j = 0$ and $\sum_{j=1}^m p_j = 1$,

$$-m\Delta_1 - (m-1) \le \gamma(1 - mp_1) \le -m\Delta_1 + (m-1). \tag{5.4}$$

Hence,

$$\gamma\Bigl(p_1 - \frac{1}{m}\Bigr) - \frac{m-1}{m} \le \Delta_1 \le \gamma\Bigl(p_1 - \frac{1}{m}\Bigr) + \frac{m-1}{m}, \tag{5.5}$$

which shows Theorem 3.10.

Suppose now that all $p_i - p_j$ are irrational. A tie between parties i and j can occur only if $(N+\gamma)p_i - (N+\gamma)p_j \in \mathbb{Z}$. If this happens for two different house sizes N_1 and N_2, then both $(N_1+\gamma)(p_i - p_j) \in \mathbb{Z}$ and $(N_2+\gamma)(p_i - p_j) \in \mathbb{Z}$, and thus by taking the difference $(N_1 - N_2)(p_i - p_j) \in \mathbb{Z}$, which is impossible if $p_i - p_j \notin \mathbb{Q}$. Hence, our assumption implies that there is a tie between parties i and j for at most one value of N, and thus at most $\binom{m}{2}$ values of N for which there is any tie. Since we consider asymptotics, we can ignore these N and assume that there are no ties. Hence we can choose α such that there are strict inequalities in (5.1), and thus there are strict inequalities in (5.2).

By (5.2) (with strict inequalities) $\lfloor (N+\gamma)(p_j - p_1)\rfloor \in \{s_j - s_1 - 1,\, s_j - s_1\}$. Let

$$I_j := s_j - s_1 - \lfloor (N+\gamma)(p_j - p_1)\rfloor. \tag{5.6}$$

Then $I_j \in \{0,1\}$ and $I_1 = 0$. We have

$$\begin{aligned}\Delta_j - \Delta_1 &= s_j - s_1 - N(p_j - p_1)\\ &= s_j - s_1 - (N+\gamma)(p_j - p_1) + \gamma(p_j - p_1)\\ &= I_j - \{(N+\gamma)(p_j - p_1)\} + \gamma(p_j - p_1).\end{aligned} \tag{5.7}$$

Summing over j we obtain, again recalling $\sum_{j=1}^m \Delta_j = 0$,

$$-m\Delta_1 = \sum_{j=1}^m I_j - \sum_{j=1}^m \{(N+\gamma)(p_j - p_1)\} + \gamma(1 - mp_1). \tag{5.8}$$

Let $L := \sum_{j=1}^m I_j$. By (5.8), we then have the formula

$$\Delta_1 = \gamma\Bigl(p_1 - \frac{1}{m}\Bigr) + \frac{1}{m}\left(\sum_{j=2}^m \{(N+\gamma)(p_j - p_1)\} - L\right). \tag{5.9}$$

Note that L is an integer with $0 \le L \le m-1$ and, by (5.6),

$$\begin{aligned}L := \sum_{j=1}^m I_j &= \sum_{j=1}^m s_j - ms_1 - \sum_{j=1}^m \lfloor (N+\gamma)(p_j - p_1)\rfloor\\ &= N - ms_1 - \sum_{j=2}^m \lfloor (N+\gamma)(p_j - p_1)\rfloor\\ &\equiv N - \sum_{j=2}^m \lfloor (N+\gamma)(p_j - p_1)\rfloor \pmod m.\end{aligned}$$

Let $\operatorname{Mod}_m(x) := m\{x/m\}$, the remainder when x is divided by m. We thus have

$$L = \operatorname{Mod}_m\Bigl(N - \sum_{j=2}^m \lfloor (N+\gamma)(p_j - p_1)\rfloor\Bigr). \tag{5.10}$$

Let $y_j := p_j - p_1$ and $a_j := \gamma(p_j - p_1)$; thus $(N+\gamma)(p_j - p_1) = Ny_j + a_j$. Let us now assume that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. It is easily seen that then $y_2, \dots, y_m$ and $\sum_1^m p_j = 1$ also are linearly independent over $\mathbb{Q}$. (This can be seen as a change of basis, using a non-singular integer matrix, in a vector space of dimension m over $\mathbb{Q}$.)

We need the following extension of Lemma 4.1, slightly generalizing [17, Lemma 4.3]. The proof is given in Appendix C.

Lemma 5.1. *Suppose that $y_1, \dots, y_k$ and 1 are linearly independent over $\mathbb{Q}$, and let $a_1, \dots, a_k$ be any real numbers.*

Let $\ell_n = \operatorname{Mod}_m\bigl(n - \sum_{j=1}^k \lfloor ny_j + a_j \rfloor\bigr) \in \{0, \dots, m-1\}$. Then the sequence of vectors

$$(\{ny_1 + a_1\}, \dots, \{ny_k + a_k\}, \ell_n) \in [0,1)^k \times \{0, \dots, m-1\} \tag{5.11}$$

is uniformly distributed in $[0,1)^k \times \{0, \dots, m-1\}$.

Note that L is (5.10) equals ℓ_N in Lemma 5.1, with y_j and a_j as chosen above. Consequently, (5.9) and Lemma 5.1 show that

$$\Delta_1 \xrightarrow{\mathrm{d}} \gamma\Bigl(p_1 - \frac{1}{m}\Bigr) + \frac{1}{m}\left(\sum_{j=2}^m U_j - \bar{L}\right), \tag{5.12}$$

where $U_j \sim \mathrm{U}(0,1)$ and $\bar{L}$ is uniformly distributed on $\{0, \dots, m-1\}$, with all variable independent. Moreover, $U_2 \overset{\mathrm{d}}{=} 1 - U_2$ and $\bar{L} + U_2 \sim \mathrm{U}(0, m)$, and thus

$$\bar{L} - U_2 \overset{\mathrm{d}}{=} \bar{L} + U_2 - 1 \overset{\mathrm{d}}{=} mU_0 - 1 \overset{\mathrm{d}}{=} m\widetilde{U}_0 + (m-2)/2.$$

Hence, (5.12) can be written

$$\Delta_1 \xrightarrow{\mathrm{d}} \gamma\Bigl(p_1 - \frac{1}{m}\Bigr) + \frac{1}{m}\left(\sum_{j=3}^m \widetilde{U}_j - m\widetilde{U}_0\right),$$

with $\widetilde{U}_j := U_j - \frac{1}{2}$ for $j \geq 1$, which by $\widetilde{U}_0 \overset{\mathrm{d}}{=} -\widetilde{U}_0$ proves Theorem 3.13. ◻

Proof of Theorem 3.11 and Corollary 3.14. Since Δ_i is bounded, e.g. by Theorem 3.10, the convergence in distribution $\Delta_i \xrightarrow{\mathrm{d}} Y_i$ in Theorem 3.13 implies convergence of all moments and it suffices to calculate the expectation and variance of Y_i. This is immediate since $\mathbb{E}\,\widetilde{U}_k = 0$ and $\operatorname{Var}\widetilde{U}_k = 1/12$; thus by (3.17),

$$\mathbb{E}\,Y_i = \gamma\Bigl(p_i - \frac{1}{m}\Bigr),$$
$$\operatorname{Var} Y_i = \frac{1}{12} + \frac{m-2}{12m^2}.$$

Theorem 3.11 and Corollary 3.14 follow. ◻

Proof of Theorem 3.16. Let $W_i := \{(N+\gamma)p_i\}$. For any j, $\{p_i\}_{i \neq j}$ and $1 = \sum_{i=1}^m p_i$ are linearly independent over $\mathbb{Q}$, and thus the vectors $(W_i)_{i \neq j}$

are uniformly distributed in $[0,1)^{m-1}$ by Lemma 4.1. Moreover,

$$\sum_{i=1}^m W_i \equiv \sum_{i=1}^m (N+\gamma)p_i = N+\gamma \equiv \gamma \pmod 1,$$

so $W_j \equiv \gamma - \sum_{i\neq j} W_i \pmod 1$ and

$$W_j = \Bigl\{\gamma - \sum_{i\neq j} W_i\Bigr\}.$$

Consequently, $(W_1,\dots,W_m) \xrightarrow{\mathrm{d}} (V_1,\dots,V_m)$ where each $V_i \sim \mathrm{U}(0,1)$, any $m-1$ of them are independent, and any such set determines the last variable by $V_j = \{\gamma - \sum_{i\neq j} V_i\}$. Obviously, the distribution of $(V_1,\dots,V_m)$ is symmetric.

We have

$$\Delta_i - \gamma p_i = s_i - (N+\gamma)p_i = s_i - \lfloor (N+\gamma)p_i \rfloor - W_i. \tag{5.13}$$

By the construction of the γ-quota method, the numbers $s_i - \lfloor (N+\gamma)p_i \rfloor$ depend only on $(W_1,\dots,W_m)$, say $s_i - \lfloor (N+\gamma)p_i \rfloor = f_i(W_1,\dots,W_m)$ for some function f_i, and it follows from (5.13) that, jointly for all i,

$$\Delta_i - \gamma p_i \xrightarrow{\mathrm{d}} Z_i := f_i(V_1,\dots,V_m) - V_i, \tag{5.14}$$

where the distribution of $(Z_1,\dots,Z_m)$ is symmetric. (The m functions f_i differ only by permutations of the variables.) Now let $\bar{Y}_i := Z_i + \gamma p_i$. Then $\Delta_i \xrightarrow{\mathrm{d}} \bar{Y}_i$, and thus $\bar{Y}_i \overset{\mathrm{d}}{=} Y_i$ by comparison with Theorem 3.13. □

Proof of Corollary 3.17. By Theorem 3.16 and the fact that each Δ_i is bounded, $\operatorname{Cov}(\Delta_i,\Delta_j) \to \operatorname{Cov}(\bar{Y}_i,\bar{Y}_j)$. Moreover, by symmetry, $\operatorname{Cov}(\bar{Y}_i,\bar{Y}_j)$ is the same for all pairs (i,j) with $i\neq j$. We have $\sum_{i=1}^m \bar{Y}_i = 0$ as a consequence of (2.3), and thus

$$0 = \operatorname{Cov}\Bigl(\bar{Y}_i, \sum_{j=1}^m \bar{Y}_j\Bigr) = \operatorname{Var}(\bar{Y}_i) + \sum_{j\neq i} \operatorname{Cov}(\bar{Y}_i,\bar{Y}_j);$$

hence by symmetry $\operatorname{Cov}(\bar{Y}_i,\bar{Y}_j) = -\operatorname{Var}(\bar{Y}_i)/(m-1) = -\operatorname{Var}(Y_i)/(m-1)$ for $i\neq j$. The final equality follows from (3.18). □

Remark 5.2. We can also see when equality holds in Theorem 3.10. The argument yielding (5.5) shows that Δ_1 is maximal exactly when there is equality in the left inequality in (5.1) for party 1, and in the right inequality for all other parties; Δ_1 is minimal when the opposite holds. It follows that, exactly as for Theorem 3.3 in Remark 4.2, Δ_i equals the upper [lower] bound in (3.15) if and only if all parties tie in (2.8) and, moreover, party i is the only lucky [unlucky] party.

6. Random party sizes

As said in the introduction, there is a long tradition of studying the bias of election methods theoretically with a fixed house size and random party sizes. More precisely, in this approach one (usually) assumes that the vector $(p_1, \dots, p_m)$ is random and uniformly distributed over the simplex $\mathfrak{S}_m := \{(p_1, \dots, p_m) \in (0,1)^m : \sum_{i=1}^m p_i = 1\}$ for some given m. One then orders the party sizes $p_1, \dots, p_m$ as $p_{[1]} \geq \dots \geq p_{[m]}$ and considers the seat excesses $\Delta_{[1]}, \dots, \Delta_{[m]}$ of the parties in the same order. (Thus, $\Delta_{[1]}$ is the seat excess of the largest party, and so on.) Results for the mean $\mathbb{E}\,\Delta_{[k]}$ of the seat excess for the k:th largest party (either exact formulas for a fixed N or asymtotic formulas as $N \to \infty$) have been obtained by Pólya [23, 24, 25], Schuster, Pukelsheim, Drton and Draper [28], Heinrich, Pukelsheim and Schwingenschlögl [13] and Drton and Schwingenschlögl [30, 6]; results on the variance are given by Schwingenschlögl and Drton [31]. (We follow these papers and order the parties in decreasing order, cf. Appendix D.)

Heinrich, Pukelsheim and Schwingenschlögl [12, 13] and Schwingenschlögl [29] have also, more generally, considered other distributions of $p_1, \dots, p_m$ (with an absolutely continuous distribution on $\mathfrak{S}_m$). Note also that Sainte-Laguë [27] considered a different distribution for two parties, see Example 3.5 (with p_2/p_1 uniformly distributed, which is not the same as p_2 uniform).

We cannot obtain these results for a given N directly from our results for random N, nor can we obtain our results for fixed $p_1, \dots, p_m$ from these results. However, by combining the methods, we can translate one type of results to the other and thus see that our main results transfer to the case of random $p_1, \dots, p_m$. (Now with random p_i in e.g. (3.7) and (3.17).)

Theorem 6.1. *Let $p_1, \dots, p_m$ be random with an absolutely continuous distribution on the simplex $\mathfrak{S}_m$, and let $N \to \infty$ (with N deterministic).*

- (i) *For the β-linear divisor method, the conclusions (3.7)–(3.10) of Theorem 3.7 hold, with $\widetilde{U}_i$ and U_i independent of $p_1, \dots, p_m$ and with $\mathbb{P}(J = j \mid p_1, \dots, p_m) = p_j$.*
- (ii) *For the γ-quota method, the conclusions of Theorems 3.13 and 3.16 hold; in particular (3.17) holds, with $\widetilde{U}_k \sim \mathrm{U}(-\frac{1}{2}, \frac{1}{2})$ independent of each other and of $p_1, \dots, p_m$.*

Proof. Consider first the β-linear divisor method. The proof of [13, Theorem 3.1(i)] shows that as $N \to \infty$,

$$(\Delta_1, \dots, \Delta_m) \xrightarrow{\mathrm{d}} (Z_1, \dots, Z_m) \tag{6.1}$$

for some random variables Z_j; with the notation in [13],

$$Z_j = -\big(U_j + q - \tfrac{1}{2} + \operatorname{sgn}(D_c) m_j(D_c)\big) + c(q - \tfrac{1}{2})W_j. \tag{6.2}$$

Here W_j is our p_j, c is our m, q is our β, U_j is our $\widetilde{U}_j$ and $\operatorname{sgn}(D_c) m_j(D_c)$ is an explicit but rather complicated function of $U_1, \dots, U_m$ and $W_1, \dots, W_m$;

however, the point is that for us it suffices to know the existence of some variables Z_j such that (6.1) holds; we can ignore the construction. The proof of (6.1) is based on the fact that for any (deterministic) sequence $\nu_n \to \infty$, the vectors of fractional parts $(\{\nu_n p_1\}, \dots, \{\nu_n p_{m-1}\})$ are uniformly distributed in $[0,1)^{m-1}$, see Lemma C.1, which is the analogue of Lemma 4.1 in the present context.

Consider now N uniformly random on $\{1, \dots, N_0\}$ and let $N_0 \to \infty$, i.e., suppose $N \xrightarrow{\mathrm{p}*} \infty$ in the notation used in the previous sections. Then $N \xrightarrow{\mathrm{p}} \infty$, so by conditioning on N, it follows that (6.1) holds for such random N as well. On the other hand, since the distribution of $p_1, \dots, p_m$ is absolutely continuous, almost surely $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. Thus, by conditioning on $p_1, \dots, p_m$ we may apply Theorem 3.7, which shows that

$$(\Delta_1, \dots, \Delta_m) \xrightarrow{\mathrm{d}} (X_1, \dots, X_m), \tag{6.3}$$

where $X_1, \dots, X_m$ are given by (3.9)–(3.10) (with our random $p_1, \dots, p_m$). In the case of random $p_1, \dots, p_m$ and random $N \xrightarrow{\mathrm{p}*} \infty$, we thus have both (6.1) and (6.3); consequently,

$$(Z_1, \dots, Z_m) \overset{\mathrm{d}}{=} (X_1, \dots, X_m). \tag{6.4}$$

Consider again deterministic $N \to \infty$. We know that (6.1) holds, so (6.4) implies that (6.3) holds in this case too. Finally, $(\bar{X}_i \mid p_i) \overset{\mathrm{d}}{=} (X_i \mid p_i)$ by Theorem 3.7, and thus $\bar{X}_i \overset{\mathrm{d}}{=} X_i$, so (3.7) too holds in the present setting.

For the γ-quota method, the proof of Theorem 3.16 applies to the present situation as well, provided we use Lemma C.1 instead of Lemma 4.1. In particular, (5.14) holds, with $V_1, \dots, V_m \sim \mathrm{U}(0,1)$ independent of $p_1, \dots, p_m$. Further. $(\bar{Y}_i \mid p_1, \dots, p_m) = (Y_i \mid p_1, \dots, p_m)$, and thus $\bar{Y}_i \overset{\mathrm{d}}{=} Y_i$ in the case of random $p_1, \dots, p_m$ too. The results follow. □

Remark 6.2. By considering $p_1, \dots, p_m$ uniformly distributed on a small neighbourhood $\max_i |p_i - p_i^0| < \varepsilon$ of some $(p_1^0, \dots, p_m^0) \in \mathfrak{S}_m$, and then letting $\varepsilon \to 0$, we recover the limit distributions in Section 3, with a somewhat different interpretation as a double limit. (In this version $p_1, \dots, p_m$ can be any numbers, including rationals.)

In particular, we can apply Theorem 6.1 to the case of uniformly distributed $p_1, \dots, p_m$ taken in decreasing order as above, since then $(p_{[1]}, \dots, p_{[m]})$ has a uniform distribution on the subset $\mathfrak{S}_{m,\geq} := \{(p_1, \dots, p_m) \in \mathfrak{S}_m : p_1 \geq \dots \geq p_m\}$ of $\mathfrak{S}_m$. This gives immediately the following theorem. (We leave the case of joint distributions to the reader.)

Theorem 6.3. *Let $p_1, \dots, p_m$ be random and uniformly distributed on $\mathfrak{S}_m$, and let $N \to \infty$.*

(i) *For the β-linear divisor method, for each $j \le m$,*

$$\Delta_{[j]} \xrightarrow{\mathrm{d}} \widehat{X}_j := (\beta - \tfrac{1}{2})(mp_{[j]} - 1) + \widetilde{U}_0 + p_{[j]} \sum_{k=1}^{m-2} \widetilde{U}_k. \tag{6.5}$$

(ii) *For the γ-quota method, for each $j \le m$,*

$$\Delta_{[j]} \xrightarrow{\mathrm{d}} \widehat{Y}_j := \gamma\Bigl(p_{[j]} - \frac{1}{m}\Bigr) + \widetilde{U}_0 + \frac{1}{m} \sum_{k=1}^{m-2} \widetilde{U}_k. \tag{6.6}$$

Here $\widetilde{U}_k \sim \mathrm{U}(-\frac{1}{2}, \frac{1}{2})$ are independent of each other and of $p_1, \dots, p_m$. □

As before, this theorem implies moment convergence because $\Delta_{[j]}$ is bounded. In particular, for the bias in this approach we obtain the following result. This was (except for the case $\gamma \ne 0$ in (ii)) conjectured by Schuster, Pukelsheim, Drton and Draper [28] and proven by Drton and Schwingenschlögl [6] and Heinrich, Pukelsheim and Schwingenschlögl [13].

Theorem 6.4. *Let $p_1, \dots, p_m$ be random and uniformly distributed on $\mathfrak{S}_m$, and let $N \to \infty$.*

(i) *For the β-linear divisor method, for each $j \le m$,*

$$\mathbb{E}\,\Delta_{[j]} \to \mathbb{E}\,\widehat{X}_j = (\beta - \tfrac{1}{2})(m\,\mathbb{E}\,p_{[j]} - 1) = (\beta - \tfrac{1}{2})\Bigl(\sum_{i=j}^{m} \frac{1}{i} - 1\Bigr). \tag{6.7}$$

(ii) *For the γ-quota method, for each $j \le m$,*

$$\mathbb{E}\,\Delta_{[j]} \to \mathbb{E}\,\widehat{Y}_j = \gamma\Bigl(\mathbb{E}\,p_{[j]} - \frac{1}{m}\Bigr) = \frac{\gamma}{m}\Bigl(\sum_{i=j}^{m} \frac{1}{i} - 1\Bigr). \tag{6.8}$$

Proof. The convergence to $\mathbb{E}\,\widehat{X}_j$ or $\mathbb{E}\,\widehat{Y}_j$ follows, as just said, from Theorem 6.3. The first equalities in (6.7) and (6.8) follow immediately from (6.5) and (6.6) since $\mathbb{E}\,\widetilde{U}_k = 0$, cf. (3.5) and (3.16). It remains only to compute $\mathbb{E}\,p_{[j]}$. This is a standard formula, and for completeness we give a proof in Appendix D, see (D.6). □

We similarly obtain results for the variance. This was found for $m = 2$ and 3 (except the case $\gamma \ne 0$) by Schwingenschlögl and Drton [31].

Theorem 6.5. *Let $p_1, \dots, p_m$ be random and uniformly distributed on $\mathfrak{S}_m$, and let $N \to \infty$.*

(i) *For the β-linear divisor method, for each $j \le m$,*

$$\begin{aligned}\operatorname{Var}\Delta_{[j]} \to \operatorname{Var}\widehat{X}_j &= \frac{1}{12}\bigl(1+(m-2)\,\mathbb{E}\,p_{[j]}^2\bigr) + (\beta-\tfrac12)^2 m^2 \operatorname{Var}(p_{[j]}) \\ &= \frac{1}{12} + \frac{m-2}{12m(m+1)}\sum_{i=j}^m \frac{1}{i^2} + \frac{m-2}{12m(m+1)}\Bigl(\sum_{i=j}^m \frac{1}{i}\Bigr)^2 \\ &\quad + (\beta-\tfrac12)^2\left(\frac{m}{m+1}\sum_{i=j}^m\frac{1}{i^2} - \frac{1}{m+1}\Bigl(\sum_{i=j}^m\frac1i\Bigr)^2\right). \end{aligned} \tag{6.9}$$

(ii) *For the γ-quota method, for each $j \le m$,*

$$\begin{aligned}\operatorname{Var}\Delta_{[j]} \to \operatorname{Var}\widehat{Y}_j &= \frac{1}{12}\Bigl(1+\frac{m-2}{m^2}\Bigr) + \gamma^2 \operatorname{Var}(p_{[j]}) \\ &= \frac{(m+2)(m-1)}{12m^2} + \gamma^2\left(\frac{1}{m(m+1)}\sum_{i=j}^m\frac{1}{i^2} - \frac{1}{m^2(m+1)}\Bigl(\sum_{i=j}^m\frac1i\Bigr)^2\right). \end{aligned} \tag{6.10}$$

Proof. Again, by Theorem 6.3, we only have to calculate $\operatorname{Var}\widehat{X}_j$ and $\operatorname{Var}\widehat{Y}_j$; the calculations are similar and we give the details only for $\widehat{X}_j$. We use the standard decomposition

$$\operatorname{Var}\widehat{X}_j = \mathbb{E}\bigl(\operatorname{Var}(\widehat{X}_j \mid p_1,\dots,p_m)\bigr) + \operatorname{Var}\bigl(\mathbb{E}(\widehat{X}_j \mid p_1,\dots,p_m)\bigr). \tag{6.11}$$

The conditional variance and mean are obtained immediately from (6.5), which yields, cf. (3.5) and (3.11),

$$\mathbb{E}(\widehat{X}_j \mid p_1,\dots,p_m) = (\beta-\tfrac12)(mp_{[j]}-1), \tag{6.12}$$

$$\operatorname{Var}(\widehat{X}_j \mid p_1,\dots,p_m) = \frac{1+(m-2)p_{[j]}^2}{12}. \tag{6.13}$$

The first equality in (6.9) follows from (6.11)–(6.13). It remains only to find $\mathbb{E}\,p_{[j]}^2$ and $\operatorname{Var}(p_{[j]})$; these are well-known and we give a computation in Appendix D for completeness, see (D.6)–(D.8). □

Example 6.6. For $m=3$ and Jefferson/D'Hondt's method we obtain from (6.7) with $\beta=1$, $\mathbb{E}(\Delta_{[1]},\Delta_{[2]},\Delta_{[3]}) \to (\frac{5}{12},-\frac{1}{12},-\frac{4}{12})$, as found already by Pólya [25].

For $m=3$ and Droop's method we obtain from (6.8) with $\gamma=1$, $\mathbb{E}(\Delta_{[1]},\Delta_{[2]},\Delta_{[3]}) \to (\frac{5}{18},-\frac{1}{18},-\frac{4}{18})$.

Example 6.7. For $m=3$, we have for the β-linear divisor method

$$(\operatorname{Var}\Delta_{[1]},\operatorname{Var}\Delta_{[2]},\operatorname{Var}\Delta_{[3]}) \to \Bigl(\frac{301}{2592}+\frac{13}{72}(\beta-\tfrac12)^2, \frac{235}{2592}+\frac{7}{72}(\beta-\tfrac12)^2, \frac{220}{2592}+\frac{4}{72}(\beta-\tfrac12)^2\Bigr)$$

and for the γ-quota method

$$(\operatorname{Var}\Delta_{[1]}, \operatorname{Var}\Delta_{[2]}, \operatorname{Var}\Delta_{[3]}) \to \Bigl(\frac{5}{54}+\frac{13}{648}\gamma^2, \frac{5}{54}+\frac{7}{648}\gamma^2, \frac{5}{54}+\frac{4}{648}\gamma^2\Bigr).$$

These were found by Schwingenschlögl and Drton [31] (where, however, the case $\gamma \neq 0$ is not treated).

Asymptotic covariances can be computed in the same way, using (3.12), (3.19) and (D.10).

In the same way as in Theorems 6.4–6.5, we can obtain results for the mean and variance of $\Delta_{[j]}$ for random $p_1, \dots, p_m$ with other distributions than the uniform one. In particular, this gives the formulas by Heinrich, Pukelsheim and Schwingenschlögl [13] and Schwingenschlögl [29] for the asymptotic bias in terms of $\mathbb{E}\, p_{[j]} = \mathbb{E}(p_j \mid p_1 \geq \dots \geq p_m)$ (assuming that $(p_1, \dots, p_m)$ has a symmetric distribution).

For example, Heinrich, Pukelsheim and Schwingenschlögl [13] consider, as a special case, the case of a threshold t (with $0 < t < 1/m$) and uniform distribution on $p_1 \geq \dots \geq p_m \geq t$. In this case, one has (as in [13]) $(p_1 - 1/m, \dots, p_m - 1/m) \overset{\mathrm{d}}{=} (1-mt)(p_1^* - 1/m, \dots, p_m^* - 1/m)$, where $(p_1^*, \dots, p_m^*)$ has a uniform distribution on $\mathfrak{S}_{m,\geq}$, and the mean and variance are easily computed.

7. The probability of violating quota

We say that a seat assignment s_i *satisfies lower quota* if $s_i \geq \lfloor q_i \rfloor$ and *satisfies upper quota* if $s_i \leq \lceil q_i \rceil$; it *satisfies quota* if both hold. In terms of the seat excess $\Delta_i = s_i - q_i$, the assignment satisfies lower [upper] quota if and only if $\Delta_i > -1$ [$\Delta_i < 1$], and it satisfies quota if and only if $|\Delta_i| < 1$.

As is well-known, Hamilton/Hare's method always satisfies quota, while Jefferson's and Droop's methods satisfy lower quota and Adams method satisfies upper quota, see e.g. [1] or Theorems 3.3 and 3.10 above. It is also well-known that Webster/Sainte-Laguë does not always satisfy quota, but that violations are unusual in practice.

Theorems 3.7 and 3.13 enable us to calculate the (asymptotic) probabilities that quota is violated for various methods. We give two examples.

Example 7.1. Consider Jefferson/D'Hondt's method. By Theorem 3.3 with $\beta = 1$, $p_i - 1 \leq \Delta_i \leq (m-1)p_i$; hence lower quota is always satisfied (as is well-known [1]), but upper quota can be violated for a party with $p_i \geq 1/(m-1)$. Consider the case $m = 3$, and suppose that $p_1 \geq 1/2$. (Then $p_2, p_3 < 1/2$, so quota is always satisfied for the other two parties; thus at most one party can violate quota.) By Theorem 3.7 with $\beta = 1$, letting $\widetilde{U}_i = \frac{1}{2} - U_i$,

$$\Delta_1 \xrightarrow{\mathrm{d}} \bar{X}_1 = \tfrac{1}{2}(3p_1 - 1) + \widetilde{U}_0 + p_1 \widetilde{U}_1 = 2p_1 - U_0 - p_1 U_1; \tag{7.1}$$

hence the asymptotic probability of violating (upper) quota is

$$\mathbb{P}(\bar{X}_1 \geq 1) = \mathbb{P}\big(U_0 + p_1 U_1 < 2p_1 - 1\big) = \frac{(2p_1 - 1)^2}{2p_1}. \tag{7.2}$$

(The set of allowed (U_0, U_1) is a right-angled triangle with sides $2p_1 - 1$ and $(2p_1 - 1)/p_1$.)

If we let (p_1, p_2, p_3) be random and uniformly distributed on $\mathfrak{S}_3$, as in Section 6, then p_1 has the density $2(1-p_1)$ and thus the asymptotic probability that party 1 violates (upper) quota is

$$\int_{1/2}^{1} \frac{(2p-1)^2}{2p} \cdot 2(1-p)\,\mathrm{d}p = \ln 2 - 2/3 \approx 0.026. \tag{7.3}$$

Consequently, the asymptotic probability that some party violates (upper) quota is

$$3 \ln 2 - 2 \approx 0.079, \tag{7.4}$$

as found by Niemeyer and Wolf [21].

Example 7.2. Consider Jefferson/D'Hondt's method, and a party i with three times the average size: $p_i = 3/m$. Then the bias $\mathbb{E}\,\bar{X}_i = 1$, by Theorem 3.4 with $\beta = 1$. It follows by (3.7) and symmetry that $\mathbb{P}(\bar{X}_i > 1) = 1/2$, so the (asymptotic) probability that the party violates quota is $1/2$. For a larger party, the probability is even greater.

Example 7.3. The Swedish parliament contains after the general election in 2010 8 parties: two large with 30% of the votes each and 6 small with 5–8% percent each. The seats are in principle distributed by Sainte-Laguë's method. (We ignore here complications due to the division into 29 constituencies and the system with adjustment seats, which in 2010 did not give complete adjustment because the number of adjustment seats was insufficient.)

We note from Theorem 3.3 (with $\beta = \frac{1}{2}$) that the small parties always satisfy quota. In fact, (3.4) shows that for Webster/Sainte-Laguë, only parties with $p_i \geq 1/(m-2)$ can violate quota.

For the large parties we have (approximatively) $p_i = 0.3$, and thus by (3.7) (still with $\beta = \frac{1}{2}$) $\bar{X}_i = \widetilde{U}_0 + 0.3 \sum_{k=1}^{6} \widetilde{U}_k$. An integration (using Maple) yields $\mathbb{P}(\bar{X}_i \geq 1) = \mathbb{P}(\bar{X}_i \leq -1) \approx 0.00045$. Hence, for each of the two large parties the (asymptotic) probability of violating quota is 0.0009.

8. Apparentements

Suppose that two or more parties decide to form a coalition in the election, so that their votes are counted together. (In some election systems, parties can register such a coalition, called an apparentement, and continue to have separate names and lists. Otherwise, the parties can always appear on election day as one party with a common umbrella name. In any case we assume that this does not attract or repel any voters, so the coalition gets exactly as many votes as the parties would have had separately, and that all other parties remain unaffected with the same number of votes as before.)

It is well-known that for a β-linear divisor method with $\beta \geq 1$ (e.g. Jefferson/D'Hondt's method) parties can never lose by forming an apparentement; they will get at least as many seats together as if they appear separately. When $\beta \leq 0$ (e.g. Adams's method), the opposite is true, and parties will never gain by forming an apparentement (conversely, they may gain by splintering), and for $0 < \beta < 1$ (e.g. Webster/Sainte-Laguë's method), they may both gain and lose. For D'Hondt's method, this effect is politically important in reality, especially in small constituencies, and apparentements are or have been a regular feature in many countries, see e.g. [4].

Leutgäb and Pukelsheim [19] give a detailed study of the resulting gains in a set of real elections, including examples, statistics and theoretical values assuming that the party sizes are random (as in Section 6). We can now give similar theoretical results for given party sizes. For simplicity, we consider only the case of a single apparentement of two parties; larger apparentements can be treated similarly as well as several apparentements (their effects are, asymptotically at least, additive).

Let the parties be i and j, and consider the expectation of the gain $s_{ij} - (s_i + s_j)$, where s_{ij} is the number of seats the parties get together as an apparentement.

Theorem 8.1. *Suppose that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$. If two parties i and j form an apparentement in the election, the mean of their seat gain $s_{ij} - s_i - s_j$ satisfies, as $N \xrightarrow{\mathrm{p*}} \infty$,*

$$\mathbb{E}(s_{ij} - s_i - s_j) \to (\beta - \tfrac{1}{2})\bigl(1 - p_i - p_j\bigr) \tag{8.1}$$

for the β-linear divisor method, and

$$\mathbb{E}(s_{ij} - s_i - s_j) \to \gamma\Bigl(\frac{2}{m} - \frac{1}{m-1}\Bigr) = \gamma\frac{m-2}{m(m-1)} \tag{8.2}$$

for the γ-quota method.

Proof. If Δ_{ij} is the seat excess for the apparentement, then

$$s_{ij} - s_i - s_j = N(p_i + p_j) + \Delta_{ij} - (Np_i + \Delta_i) - (Np_j + \Delta_j) = \Delta_{ij} - \Delta_i - \Delta_j,$$

so $\mathbb{E}(s_{ij} - s_i - s_j) = \mathbb{E}\,\Delta_{ij} - \mathbb{E}\,\Delta_i - \mathbb{E}\,\Delta_j$, which by Theorems 3.4 and 3.11 converges to, recalling that Δ_{ij} is the seat excess in a contest with $m-1$ parties,

$$(\beta - \tfrac{1}{2})\Bigl((m-1)(p_i + p_j) - 1 - (mp_i - 1) - (mp_j - 1)\Bigr)$$

and

$$\gamma\Bigl(\Bigl(p_i + p_j - \frac{1}{m-1}\Bigr) - \Bigl(p_i - \frac{1}{m}\Bigr) - \Bigl(p_j - \frac{1}{m}\Bigr)\Bigr)$$

for the β-linear divisor method and γ-quota method, respectively. $\square$

One sees in the same way that these expected gains (which are negative if $\beta < 1/2$ or $\gamma < 0$) are balanced by a loss for each other party k of $(\beta - \frac{1}{2})p_k$ for the β-linear divisor method and $\gamma/(m(m-1))$ for the γ-quota method.

We have seen that apparentements are favoured when $\beta > 1/2$ or $\gamma > 0$, but the effect is rather small, especially for quota methods. (Nevertheless, as said above, the effect is large enough to be politically important in reality for D'Hondt's method.) Note that for the β-linear divisor method, the gain in (8.1) does not depend on the number (or sizes) ot the other parties, while for the γ-quota method, the gain in (8.2) is largest for $m = 3$ or 4 and decreases as $O(1/m)$ for large m.

Example 8.2. For Jefferson/D'Hondt's method, two small parties that form an apparentement gain at most 0.5 seats together, and two large parties gain less.

Example 8.3. For Droop's method, any two parties that form an apparentement gain at most $\frac{1}{6} \approx 0.167$ seats together when there are 3 or 4 parties, 0.15 seats when there are 5 parties, and less if there are more.

Problem 8.4. What are, for the different methods, the asymptotic distributions of the seat gain $s_{ij} - s_i - s_j$ for an apparentement?

We have so far studied the gain of the two parties combined, but the parties are probably more interested in how the gain is split between them. Typically, the seats given to an apparentement are distributed, in a *sub-apportionment*, between the participating parties by the same election method as for the main distribution (the *super-apportionment*). It seems reasonable that the gain then, on the average, is split between the parties proportionally to their sizes, cf. Theorems A.11 and B.3. However, we shall see that this holds only for the divisor methods.

Consider first the β-linear divisor method. Theorem 3.4 shows that, asymptotically as $N \xrightarrow{\text{p*}} \infty$, the number of seats the apparentement gets is on the average

$$N(p_i + p_j) + (\beta - \tfrac{1}{2})\big((m-1)(p_i + p_j) - 1\big). \tag{8.3}$$

If we use Theorem 3.4 again for the distribution of these seats between i and j, we obtain that i gets on the average

$$\Big(N(p_i + p_j) + (\beta - \tfrac{1}{2})\big((m-1)(p_i + p_j) - 1\big)\Big)\frac{p_i}{p_i + p_j} + (\beta - \tfrac{1}{2})\Big(2\frac{p_i}{p_i + p_j} - 1\Big)$$
$$= Np_i + (\beta - \tfrac{1}{2})\Big((m-1)p_i + \frac{p_i}{p_i + p_j} - 1\Big) \tag{8.4}$$

seats, which, comparing to Theorem 3.4 for an election without the apparentement, means an average gain for party i of

$$(\beta - \tfrac{1}{2})\Big(\frac{p_i}{p_i + p_j} - p_i\Big) = \frac{p_i}{p_i + p_j}(\beta - \tfrac{1}{2})(1 - p_i - p_j), \tag{8.5}$$

which indeed is the proportional share $p_i/(p_i + p_j)$ of the joint gain in (8.1).

There is, however, a gap in this argument. In our model, the seat assignment is a deterministic function of N both in the super-apportionment and in the sub-apportionment, and the calculation above using expectations for random N assumes that there is no hidden correlation, where the numbers of seats given to the apparentement are seat numbers that tend to favour one of the two parties in the sub-apportionment. It seems very unlikely that there will be such a correlation under our assumption that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$, but we have so far no rigorous proof, so (8.5) should only be regarded as a conjecture for the asymptotic average gain for party i.

Problem 8.5. Verify rigorously that the expected gain for party i converges to the value in (8.5) as $N \xrightarrow{\text{p*}} \infty$.

Nevertheless, if we tentatively use (8.5), and regard it as a valid approximation for real elections with finite N (and arbitrary $p_1, \dots, p_m$), we can draw some practical conclusions for D'Hondt's method ($\beta = 1$). If the parties i and j both are small, then (8.5) shows that the gain for party i is almost $p_i/(p_i + p_j)$. Thus most of the gain of the apparentement goes to the larger party, and a party has very little to gain by an apparentement with a party that is substantially larger. In fact, if party i can choose between several parties to form an apparentement, we see, perhaps surprisingly, that it gets the largest gain by choosing the smallest partner. (This is based on asymptotics as $N \to \infty$. For a real election with finite N there is a trade-off since a really small partner will not help significantly; the advantage can hardly be larger than Np_j, so presumably only parties with $q_j = Np_j$ being at least 1/2 or so will be useful partners. It would be interesting to study such effects for finite N, but that is beyond the scope of the present paper.)

Consider now instead the γ-quota method. One important complication is that quota methods are not uniform, in the sense of Appendix A.3. Thus, even if the apparentement gets the same number of seats as the two parties would have got together if they had appeared separately in the election, the sub-apportionment may give a different distribution of these seats between the parties.

We calculate as above, with the same reservation that a rigorous verification still is lacking. Theorem 3.11 shows that, asymptotically as $N \xrightarrow{\text{p*}} \infty$, the number of seats the apparentement gets is on the average

$$N(p_i + p_j) + \gamma\Big((p_i + p_j) - \frac{1}{m-1}\Big). \tag{8.6}$$

If we use Theorem 3.11 again for the distribution of these seats between i and j, we obtain that i gets on the average

$$\begin{aligned}\Big(N(p_i + p_j) + \gamma\Big((p_i + p_j) - \frac{1}{m-1}\Big)\Big)\frac{p_i}{p_i + p_j} + \gamma\Big(\frac{p_i}{p_i + p_j} - \frac{1}{2}\Big) \\ = Np_i + \gamma\Big(p_i + \frac{m-2}{m-1}\frac{p_i}{p_i + p_j} - \frac{1}{2}\Big)\end{aligned} \tag{8.7}$$

seats, which, compared to Theorem 3.11 for an election without the apparentement, means an average gain for party i of

$$\gamma\Bigl(\frac{m-2}{m-1}\frac{p_i}{p_i+p_j}-\frac{m-2}{2m}\Bigr). \tag{8.8}$$

This is quite different from the proportional share $p_i/(p_i+p_j)$ of the joint gain in (8.1). In fact, we see that for $\gamma>0$ (e.g. Droop's method), a party gains only if $p_i/(p_i+p_j)>(m-1)/(2m)$. In other words, a party will *lose* if it forms an apparentement with a party that is only a little larger. Typically, the redistribution within the apparentement by the sub-apportionment is more important than the collective gain.

9. The Sainte-Laguë divergence

Sainte-Laguë [26, 27] based his proposal of his method on the fact that it is the least squares method minimising the sum

$$S:=\sum_{i=1}^m v_i\Bigl(\frac{s_i}{v_i}-\frac{N}{V}\Bigr)^2=\sum_{i=1}^m\frac{(s_i-Np_i)^2}{v_i}=\sum_{i=1}^m\frac{\Delta_i^2}{v_i}; \tag{9.1}$$

multiplying this by V we obtain the equivalent quantity

$$\tilde S:=VS=\sum_{i=1}^m\frac{\Delta_i^2}{p_i}, \tag{9.2}$$

which is called the *Sainte-Laguë divergence* by Heinrich, Pukelsheim and Schwingenschlögl [12, 13], where the asymptotic distribution is studied under the assumption of random party sizes (see Section 6). We obtain corresponding result in our setting as corollaries of the results in Section 3. (Unfortunately, we do not get as explicit result for quota methods as for divisor methods because Theorem 3.16 is less explicit.) The most interesting case is the $\frac12$-linear divisor method (Sainte-Laguë's method) [12], since then S and $\tilde S$ are the minima over all allocations, but it is also interesting to see how much larger $\tilde S$ is for other methods [13].

Theorem 9.1. *Suppose that $p_1,\dots,p_m$ are linearly independent over $\mathbb{Q}$, and let $N\xrightarrow{\mathrm{p*}}\infty$.*

(i) *For the β-linear divisor method,*

$$\tilde S\xrightarrow{\mathrm{d}}S_{\beta,m}:=\sum_{i=1}^m\frac{(V_i+\beta-1)^2}{p_i}-\Bigl(\sum_{i=1}^m(V_i+\beta-1)\Bigr)^2, \tag{9.3}$$

where $V_1,\dots,V_m$ are as in Theorem 3.7.

(ii) *For the γ-quota method,*

$$\tilde S\xrightarrow{\mathrm{d}}\bar S_{\gamma,m}:=\sum_{i=1}^m\frac{\bar Y_i^2}{p_i}, \tag{9.4}$$

where $\bar Y_1,\dots,\bar Y_m$ are as in Theorem 3.16.

Proof. For the β-linear divisor method, we use Theorem 3.7 and obtain by the continuous mapping theorem [2, Theorem 5.1] and (9.2) immediately $\tilde{S} \xrightarrow{\mathrm{d}} \sum_{i=1}^m X_i^2/p_i$. We write (3.10) as $X_i = p_i A - A_i$, where $A_i := V_i + \beta - 1$ and $A := \sum_{i=1}^m (V_i + \beta - 1) = \sum_{i=1}^m A_i$. Thus

$$\begin{aligned}\sum_{i=1}^m \frac{X_i^2}{p_i} &= \sum_{i=1}^m \frac{p_i^2 A^2 - 2p_i A A_i + A_i^2}{p_i} = \sum_{i=1}^m p_i A^2 - 2\sum_{i=1}^m A_i A + \sum_{i=1}^m \frac{A_i^2}{p_i} \\ &= \sum_{i=1}^m \frac{A_i^2}{p_i} - A^2,\end{aligned}$$

which is (9.3).

The result for the γ-quota method follows directly from Theorem 3.16. □

The expression in (9.3) for $S_{\beta,m}$ may be compared to the somewhat different expression in [12, 13] for the limit for fixed N and random $p_1, \dots, p_m$; by the argument in Section 6, the formulas have to be equivalent (for random $p_1, \dots, p_m$).

For the expectation we similarly obtain:

Theorem 9.2. *Suppose that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$, and let $N \xrightarrow{\mathrm{p*}} \infty$.*

(i) *For the β-linear divisor method,*

$$\mathbb{E}\,\tilde{S} \to \frac{1}{12}\left(\sum_{i=1}^m \frac{1}{p_i} + m - 2\right) + (\beta - \tfrac{1}{2})^2 \left(\sum_{i=1}^m \frac{1}{p_i} - m^2\right). \tag{9.5}$$

(ii) *For the γ-quota method,*

$$\mathbb{E}\,\tilde{S} \to \frac{(m+2)(m-1)}{12m^2}\sum_{i=1}^m \frac{1}{p_i} + \frac{\gamma^2}{m^2}\left(\sum_{i=1}^m \frac{1}{p_i} - m^2\right). \tag{9.6}$$

Proof. This can be shown from Theorem 9.1, but it is easier to use

$$\mathbb{E}\,\tilde{S} = \sum_{i=1}^m \frac{\mathbb{E}\,\Delta_i^2}{p_i} = \sum_{i=1}^m \frac{\operatorname{Var}\Delta_i + (\mathbb{E}\,\Delta_i)^2}{p_i}$$

and Theorem 3.4, Corollary 3.8, Theorem 3.11, Corollary 3.14. □

Note that $\sum_{i=1}^m 1/p_i \geq m^2$, with equality if and only if all p_i are equal. We know that $\tilde{S}$ always is minimised by the $\frac{1}{2}$-linear divisor method (Sainte-Laguë's method). We see from Theorem 9.2 that also asymptotically, $\mathbb{E}\,\tilde{S}$ is strictly larger for any β-linear divisor method or γ-quota method, except when all parties have exactly the same size (a trivial case where all methods yield the same result; note that we have excluded this case by our assumptions).

Furthermore, Theorems 9.1–9.2 show that $\tilde{S}$ is typically of the order m^2 if all p_i are of roughly the same order, but if some p_i is very small, $\tilde{S}$ can be substantially larger. Heinrich, Pukelsheim and Schwingenschlögl [12, 13] consider $\tilde{S}$ and the limit $S_{\beta,m}$ for random, uniformly distributed, $p_1, \dots, p_m$, and show that, as $m \to \infty$,

$$S_{\beta,m}/m^2 - \bigl(\tfrac{1}{12} + (\beta - \tfrac12)^2\bigr) \log m \xrightarrow{\mathrm{d}} S^*, \tag{9.7}$$

where S^* is a certain 1-stable random variable (depending on β). We see that the $\log m$ term and the large tail of S^* (we have $\mathbb{E}\, S^* = \infty$) depend on the possibility of some very small p_i; the asymptotic behaviour for deterministic $p_1, \dots, p_m$ with, say, $(\min p_i)^{-1} = O(m)$ is quite different. To see this better, we first replace V_i (which are dependent) by U_i (which are i.i.d.) in (9.3); the following lemma shows that for large m, the difference is small. (Cf. similar approximations in [12, 13].)

Lemma 9.3. *Let $\hat{U}_i := U_i + \beta - 1$, so $\hat{U}_i$ are i.i.d.* $\mathrm{U}(\beta - 1, \beta)$*, and let*

$$S'_{\beta,m} := \sum_{i=1}^{m} \frac{\hat{U}_i^2}{p_i} - \Bigl(\sum_{i=1}^{m} \hat{U}_i\Bigr)^2. \tag{9.8}$$

Then $\mathbb{E}\bigl|S_{\beta,m} - S'_{\beta,m}\bigr| = O(m)$, uniformly in $p_1, \dots, p_m$.

Proof. We have $|U_i+\beta-1| \le |\beta|+1 = O(1)$ and similarly $|V_i+\beta-1| = O(1)$. Further, by (3.9), $\mathbb{P}(V_i \neq U_i) = \mathbb{P}(J = i) = p_i$. Hence,

$$\mathbb{E}\, \frac{\bigl|(V_i + \beta - 1)^2 - (U_i + \beta - 1)^2\bigr|}{p_i} = \frac{O(p_i)}{p_i} = O(1). \tag{9.9}$$

Similarly, $\sum_{i=1}^m (V_i + \beta - 1) - \sum_{i=1}^m (U_i + \beta - 1) = -U_J = O(1)$ and thus

$$\Bigl(\sum_{i=1}^{m} (V_i + \beta - 1)\Bigr)^2 - \Bigl(\sum_{i=1}^{m} (U_i + \beta - 1)\Bigr)^2 = O(m). \tag{9.10}$$

The result follows by summing (9.9) and (9.10). ∎

We can now state a normal limit theorem corresponding to (9.7), but for deterministic $p_1, \dots, p_m$.

Theorem 9.4. *Suppose that for each m, we are given $p_1, \dots, p_m$ such that $m^{3/2} \min_i p_i \to \infty$ as $m \to \infty$. Let $A_m := \sum_{i=1}^m p_i^{-1}$, $B_m := \sum_{i=1}^m p_i^{-2}$ and*

$$C_m := \sum_{i=1}^{m} (p_i^{-1} - m)^2 = B_m - 2mA_m + m^3.$$

Then, as $m \to \infty$,

$$\frac{S_{\beta,m} - \bigl(A_m/12 + (A_m - m^2)(\beta - 1/2)^2\bigr)}{\sqrt{B_m/180 + (\beta - 1/2)^2 C_m/3}} \xrightarrow{\mathrm{d}} N\bigl(0, 1\bigr).$$

Proof. Note first that, e.g. by Jensen's inequality, $A_m \ge m^2$, $B_m \ge m^3$ and $C_m \ge 0$.

Write, for convenience, $b := \beta - 1/2$. We have $\hat{U}_i = \widetilde{U}_i + b$ and $\mathbb{E}\,\hat{U} = b$. Thus, (9.8) can be expanded as

$$S'_{\beta,m} = \sum_{i=1}^{m} \frac{\widetilde{U}_i^2 + 2b\widetilde{U}_i + b^2}{p_i} - \Bigl(mb + \sum_{i=1}^{m} \widetilde{U}_i\Bigr)^2$$
$$= \sum_{i=1}^{m} \Bigl(\frac{\widetilde{U}_i^2 + 2b\widetilde{U}_i}{p_i} - 2mb\widetilde{U}_i\Bigr) + b^2 A_m - b^2 m^2 - \Bigl(\sum_{i=1}^{m} \widetilde{U}_i\Bigr)^2.$$

The last term has expectation $\operatorname{Var} \sum_{i=1}^{m} \widetilde{U}_i = O(m)$, and can thus be ignored. Define

$$Z_i := \frac{\widetilde{U}_i^2 + 2b\widetilde{U}_i}{p_i} - 2mb\widetilde{U}_i.$$

Then

$$\mathbb{E}\, Z_i = \frac{\mathbb{E}\,\widetilde{U}_i^2}{p_i} = \frac{1}{12 p_i}$$

and, by a straightforward calculation,

$$\operatorname{Var} Z_i = \frac{1}{180 p_i^2} + \frac{b^2}{3}\Bigl(\frac{1}{p_i} - m\Bigr)^2.$$

The random variables Z_i are independent and each is bounded with $|Z_i| = O(1/p_i + m) = o(m^{3/2}) = o(B_m^{1/2})$. The result now follows by the central limit theorem, using either Lindeberg's or Lyapounov's condition, see e.g. [11, Section 7.2]. □

Remark 9.5. Sainte-Laguë [27] also studied the least squares functional calculated per seat and not per voter:

$$\sum_{i=1}^{m} s_i \Bigl(\frac{v_i}{s_i} - \frac{V}{N}\Bigr)^2 = \frac{V^2}{N^2} \sum_{i=1}^{m} \frac{\Delta_i^2}{s_i}, \tag{9.11}$$

and found that this is minimised by the divisor method with $d(n) = \sqrt{n(n-1)}$, later suggested by Huntington, see Appendix A and Remark A.10. The asymptotics of this quantity is, after division by $(V/N)^3$, the same as for S, because $s_i/N \to p_i$.

10. Some other goodness-of-fit functionals

Many election methods, including the ones treated here, can be characterised as minimising various functionals, see e.g. [1], [18] and [20]. Theorems 3.7 and 3.16 yield by the continuous mapping theorem asymptotic distributions for many such functionals for the methods treated here, under our standard assumptions that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$ and $N \xrightarrow{\text{p*}} \infty$, which we assume below. We give some examples, leaving others to the reader. Note, however, that our results for quota methods are less complete, cf. Problem 3.18.

10.1. **Hamilton divergences.** Hamilton/Hare's method, the 0-quota method, is the unique method minimising $\max_i |\Delta_i|$. It also the method minimising $\max_i \Delta_i$. Moreover, it is, for any convex function $\varphi : \mathbb{R} \to [0,\infty]$ with $\varphi(0) = 0$ and $\varphi(x) > 0$ for $x \neq 0$, the unique method minimising $\sum_i \varphi(\Delta_i)$, as was shown by Pólya [24].

For the β-linear divisor method, Theorem 3.7 implies

$$\max_{1\le i\le m} |\Delta_i| \xrightarrow{\mathrm{d}} \max_{1\le i\le m} |X_i| \tag{10.1}$$

with X_i given by (3.10). However, we do not know any really simple description of this limit variable.

For the γ-quota method, including Hamilton/Hare's method and thus the minimum of $\max_i |\Delta_i|$ over all allocations, Theorem 3.16 yields

$$\max_{1\le i\le m} |\Delta_i| \xrightarrow{\mathrm{d}} \max_{1\le i\le m} |Y_i|, \tag{10.2}$$

but we have no explicit description of this limit variable.

Similar results hold for $\max_i \Delta_i$.

We obtain more complete results if we instead consider the least squares functional $\sum_{i=1}^m \Delta_i^2$, which also yields Hamilton/Hare's method by Polya's theorem [24]. (This case was shown earlier by Sainte-Laguë [27], who attributed the result to Zivy.)

Theorem 10.1. *With assumptions and notations as in Section 3, for the β-linear divisor method,*

$$\sum_{i=1}^m \Delta_i^2 \xrightarrow{\mathrm{d}} S^{\mathrm{H}}_{\beta,m} := \sum_{i=1}^m X_i^2, \tag{10.3}$$

$$\mathbb{E}\sum_{i=1}^m \Delta_i^2 \to \mathbb{E}\, S^{\mathrm{H}}_{\beta,m} = \sum_{i=1}^m \frac{1+(m-2)p_i^2}{12} + (\beta-\frac12)^2\sum_{i=1}^m (mp_i-1)^2 \tag{10.4}$$

$$= \frac{(m+2)(m-1)}{12m} + \Bigl(\frac{m-2}{12} + m^2(\beta-\frac12)^2\Bigr)\sum_{i=1}^m\Bigl(p_i-\frac1m\Bigr)^2, \tag{10.5}$$

and for the γ-quota method,

$$\sum_{i=1}^m \Delta_i^2 \xrightarrow{\mathrm{d}} \bar S^{\mathrm{H}}_{\gamma,m} := \sum_{i=1}^m \bar Y_i^2, \tag{10.6}$$

$$\mathbb{E}\sum_{i=1}^m \Delta_i^2 \to \mathbb{E}\, \bar S^{\mathrm{H}}_{\gamma,m} = \frac{(m+2)(m-1)}{12m} + \gamma^2\sum_{i=1}^m\Bigl(p_i-\frac1m\Bigr)^2. \tag{10.7}$$

Proof. By Theorems 3.7 and 3.16, calculating $\mathbb{E}\, X_i^2$ and $\mathbb{E}\, \bar Y_i^2$ by Theorem 3.4, Corollary 3.8, Theorem 3.11, Corollary 3.14 as in the proof of Theorem 9.2. ∎

10.2. **Jefferson and Adams divergences.** Heinrich and Schwingenschlögl [14] consider the functionals

$$S^{\mathrm{J}} := \max_{1\le i\le m}\Bigl(\frac{s_i}{p_i}-N\Bigr) = \max_{1\le i\le m}\frac{\Delta_i}{p_i} \tag{10.8}$$

and

$$S^{\mathrm{A}} := \min_{1\le i\le m}\Bigl(\frac{s_i}{p_i}-N\Bigr) = \min_{1\le i\le m}\frac{\Delta_i}{p_i}. \tag{10.9}$$

(We could as well use the maximum and minimum of $s_i/v_i - N/V$.) Since $\sum_{i=1}^m \Delta_i = 0$, we have $S^{\mathrm{J}} \ge 0$ but $S^{\mathrm{A}} \le 0$, and we will therefore use $|S^{\mathrm{A}}| = -S^{\mathrm{A}}$.

It is easy to see that S^{J} and $|S^{\mathrm{A}}|$ are minimised, over all allocations of N seats, by the Jefferson and Adams methods, respectively [27], [1].

Theorem 3.7 yields the asymptotic distribution of S^{J} and $|S^{\mathrm{A}}|$ for any β-linear divisor method; we consider for simplicity only the optimal methods.

Theorem 10.2. *With assumptions and notations as in Section 3, for the optimal allocations,*

$$S^{\mathrm{J}} \xrightarrow{\mathrm{d}} S^{\mathrm{J}}_m := \sum_{j=1}^{m-1} U_j, \tag{10.10}$$

and

$$|S^{\mathrm{A}}| \xrightarrow{\mathrm{d}} |S^{\mathrm{A}}_m| := m - \sum_{j=1}^{m} V_j - \min_{1\le i\le m}\frac{1-V_i}{p_i}. \tag{10.11}$$

We have

$$\mathbb{E}\, S^{\mathrm{J}}_m = \mathbb{E}\,|S^{\mathrm{A}}_m| = \frac{m-1}{2}. \tag{10.12}$$

Proof. S^{J} is minimised by Jefferson's method, so Theorem 3.7 with $\beta = 1$ yields, noting that $\min V_i/p_i = 0$ by (3.9),

$$\begin{aligned} S^{\mathrm{J}} &:= \max_{1\le i\le m}\frac{\Delta_i}{p_i} \xrightarrow{\mathrm{d}} \max_{1\le i\le m}\frac{X_i}{p_i} = \sum_{j=1}^{m} V_j - \min_{1\le i\le m}\frac{V_i}{p_i} \\ &= \sum_{j=1}^{m} V_j = \sum_{j\ne J} U_j \overset{\mathrm{d}}{=} \sum_{j=1}^{m-1} U_j. \end{aligned}$$

Similarly, $|S^{\mathrm{A}}|$ is minimised by Adams's method, so Theorem 3.7 with $\beta = 0$ yields

$$|S^{\mathrm{A}}| := -\min_{1\le i\le m}\frac{\Delta_i}{p_i} \xrightarrow{\mathrm{d}} -\min_{1\le i\le m}\frac{X_i}{p_i} = -\sum_{j=1}^{m} V_j + m - \min_{1\le i\le m}\frac{1-V_i}{p_i}.$$

We have $\mathbb{E}\, S^{\mathrm{J}}_m = (m-1)/2$ directly from (10.10).

For $|S^{\mathrm{A}}_m|$, we first note that if $B := \min_{1\le i\le m} 1/p_i$, then for $0<x<B$,

$$\begin{aligned}\mathbb{P}\Bigl(\min_{1\le i\le m}\frac{1-V_i}{p_i}>x\Bigr)&=\mathbb{P}\bigl(V_i<1-xp_i,\ i=1,\dots,m\bigr)\\&=\sum_{j=1}^m p_j\,\mathbb{P}\bigl(V_i<1-xp_i,\ i=1,\dots,m\mid J=j\bigr)\\&=\sum_{j=1}^m p_j\,\mathbb{P}\bigl(U_i<1-xp_i,\ i\neq j\bigr)\\&=\sum_{j=1}^m p_j\prod_{i\neq j}(1-xp_i),\end{aligned}$$

while $\mathbb{P}\Bigl(\min_{1\le i\le m}(1-V_i)/p_i>x\Bigr)=0$ for $x\ge B$. Hence,

$$\begin{aligned}\mathbb{E}\min_{1\le i\le m}\frac{1-V_i}{p_i}&=\int_0^\infty\mathbb{P}\Bigl(\min_{1\le i\le m}\frac{1-V_i}{p_i}>x\Bigr)\,\mathrm{d}x\\&=\int_0^B\sum_{j=1}^m p_j\prod_{i\neq j}(1-xp_i)\,\mathrm{d}x=\Bigl[-\prod_{i=1}^m(1-xp_i)\Bigr]_{x=0}^B=1.\end{aligned}\tag{10.13}$$

Furthermore, $\mathbb{E}\sum_{j=1}^m V_j=\sum_{j=1}^{m-1}U_j=(m-1)/2$, and (10.12) follows from (10.11). □

Remark 10.3. Note that S^{J}_m in (10.10) does not depend on $p_1,\dots,p_m$, but for $|S^{\mathrm{A}}_m|$, the situation is more complicated; the representation in (10.11) depends on $p_1,\dots,p_m$, but it is still possible that $|S^{\mathrm{A}}_m|$ has the same distribution for all $p_1,\dots,p_m$ and that we just have found an unnecessarily complicated description of it. Note that $\mathbb{E}|S^{\mathrm{A}}_m|$ does not depend on $p_1,\dots,p_m$. In the case $m=2$, it is easy to see that $|S^{\mathrm{A}}_m|\overset{\mathrm{d}}{=}U\overset{\mathrm{d}}{=}S^{\mathrm{J}}_m$ for every (p_1,p_2), but we do not know whether something similar is true for $m\ge3$.

Problem 10.4. Is $|S^{\mathrm{A}}_m|\overset{\mathrm{d}}{=}S^{\mathrm{J}}_m$ for every $p_1,\dots,p_m$?

The asymptotic results by [14], taking the limit of the limits as $m\to\infty$, follow easily in our setting too.

Theorem 10.5. *For any $p_1,\dots,p_m$ depending on m, as $m\to\infty$,*

$$\begin{aligned}\sqrt{m}(S^{\mathrm{J}}_m-m/2)&\overset{\mathrm{d}}{\longrightarrow}N(0,\tfrac{1}{12}),\\\sqrt{m}(|S^{\mathrm{A}}_m|-m/2)&\overset{\mathrm{d}}{\longrightarrow}N(0,\tfrac{1}{12}).\end{aligned}$$

Proof. The result for S^{J}_m is immediate from (10.10) and the central limit theorem.

For $|S_m^{\mathrm{A}}|$ we use (10.11). We have $\sum_{j=1}^m (U_j - V_j) = U_J$ and consequently, by (10.11) and (10.13),

$$\mathbb{E}\left| |S_m^{\mathrm{A}}| - \sum_{j=1}^m (1-U_j) \right| \le \mathbb{E}\left| \sum_{j=1}^m (U_j - V_j) - \min_{1\le i\le m} \frac{1-V_i}{p_i} \right| \le \mathbb{E}\, U_J + \mathbb{E} \min_{1\le i\le m} \frac{1-V_i}{p_i} < 2,$$

and the result follows by the central limit theorem applied to $\sum_{j=1}^m (1 - U_j)$. □

11. Rational p_i

In our main results we have assumed that $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$, since this is necessary for Lemma 4.1. We consider briefly what happens when this is not satisfied, i.e., when $p_1, \dots, p_m$ are linearly dependent over $\mathbb{Q}$. In particular, we are for obvious practical reasons interested in the case when $p_1, \dots, p_m$ are rational. (See also the corresponding discussion in [17].)

Note first that it may now happen that there are ties. We assume that any ties are resolved randomly; thus the quantities $s_i(N)$ and $\Delta_i(N)$ in general may be random variables, also for a fixed N. This is no real problem, however, and when discussing the means we may replace them by their average over all solutions in case of a tie.

Lemma 4.1 does not apply when $p_1, \dots, p_m$ are linearly dependent over $\mathbb{Q}$, but the proof of it sketched in Section 4 shows that the sequence $(\{ny_1 + a_1\}, \dots, \{ny_k + a_k\})_{n\ge1}$ is uniformly distributed on a coset of a subgroup of $[0,1)^k$; more precisely, if we for simplicity take all $a_i = 0$, the empirical distributions converge to the uniform probability measure $\mu_{\mathbf{y}}$ on a subgroup, with Fourier coefficients given by

$$\widehat{\mu}_{\mathbf{y}}(\ell_1, \dots, \ell_k) = \begin{cases} 1 & \text{if } \sum_{j=1}^k \ell_j y_j \in \mathbb{Z}, \\ 0 & \text{otherwise.} \end{cases} \tag{11.1}$$

The proofs of Theorems 3.7 and 3.16 now show, with minor modifications, that the seat excesses Δ_i still converge jointly in distribution as $N \xrightarrow{\mathrm{p}*} \infty$, but the limit distribution is different, and depends on $p_1, \dots, p_m$ through some measures of the type (11.1). In particular, the bias $\lim \mathbb{E}\,\Delta_i$ is well-defined, but we do not have an explicit formula for it.

However, if we consider a sequence of party size distributions $(p_{1k}, \dots, p_{mk})$, $k = 1, 2, \dots$, (with a fixed m), such that $p_{ik} \to p_i$ for each i as $k \to \infty$ and further, for every integer vector $(a_1, \dots, a_m) \ne 0$, we have $\sum_{i=1}^m a_i p_{ik} \ne 0$ for all large k, then it follows from (11.1) and an inspection of the proofs of Theorems 3.7 and 3.16 that the measures $\mu_{\mathbf{y}}$ that now appear in the proofs will converge to the uniform measure μ, and thus the limit variables X_i and

$\bar{Y}_i$ will converge to the corresponding variables given in Section 3 for the case when $p_1, \dots, p_m$ are linearly independent over $\mathbb{Q}$.

This shows that the results above are good approximations also for linearly dependent $p_1, \dots, p_m$, as long as there is no linear relation (3.2) with small integers a_i. In particular, for rational $p_1, \dots, p_m$ we typically have a good approximation unless some p_i has a small denominator.

In the remainder of this section, we suppose that $p_1, \dots, p_m$ are rational, with a common denominator L. Theorems A.11 and B.3 then show that the sequence of seat excesses $\Delta_i := s_i(N) - Np_i$, $N \geq 1$, has period L. (If $\beta < 0$, $\beta > 1$ or $\gamma < 0$ we may have to except some small N.) This gives another (simpler) proof that the limit distribution of Δ_i as $N \overset{\mathrm{p*}}{\longrightarrow} \infty$ exists in this case, and shows that it is the same as the distribution if we take N uniformly in a period $\{K+1, \dots, K+L\}$ (for any K that is large enough). In particular, the asymptotic distribution is discrete, and therefore not the same as in Section 3. The asymptotic bias $\lim \mathbb{E}\,\Delta_i$ is easily calculated by taking the average of Δ_i over a period.

Example 11.1. Let $m = 2$ and take $(p_1, p_2) = (\frac{2}{3}, \frac{1}{3})$. Then the β-linear divisor method equals the γ-quota method with $\gamma = 2\beta - 1$, see Appendix B.1. Assume, for simplicity, $0 < \beta < 1$, i.e., $-1 < \gamma < 1$. A simple calculation shows that for $N = 1, 2, 3$, we have $\Delta_1 = \frac{1}{3}, -\frac{1}{3}, 0$. Hence the average is 0, so the methods are unbiased in this case, for any $\beta \in (0, 1)$ or $\gamma \in (-1, 1)$. (In fact, for this example the methods all coincide with Hamilton/Hare = Webster/Sainte-Laguë.) In particular, Theorems 3.4 and 3.11 do not hold for rational $p_1, \dots, p_m$ and $\beta \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$ and $\gamma \in (-1, 0) \cup (0, 1)$.

If we instead take $\beta = 1$, i.e. $\gamma = 1$, we have a tie for $N = 2$, with $\Delta_1 = -\frac{1}{3}$ or $\frac{2}{3}$. The average of Δ_1 over a period now is $\frac{1}{6}$.

For the methods with $\beta = 1/2$ and $\gamma = 0$, the result 0 in Example 11.1 agrees with Theorems 3.4 and 3.11. In fact, this is true for any rational $p_1, \dots, p_m$.

Theorem 11.2. *The $\frac{1}{2}$-linear divisor method (Webster/Sainte-Laguë) and the 0-quota method (largest remainder/Hamilton/Hare) are asymptotically unbiased for any rational $p_1, \dots, p_m$.*

Proof. Let, as above, L be a common denominator of $p_1, \dots, p_m$. It is for both methods easily seen, arguing similarly to the proofs of Theorems A.11 and B.3, that $s_i(L-N) = Lp_i - s_i(N)$ for $N = 0, \dots, L$. Hence, $\Delta_i(L-N) = -\Delta_i(N)$, and the average over a period $N = 1, \dots, L$ vanishes. □

Also for $\beta = 1$ and $\gamma = 1$, the result in Example 11.1 agrees with Theorems 3.4 and 3.11. For divisor methods, we have the following general result. (By Example 11.1 and (3.14), the result does not hold for any other β.)

Theorem 11.3. *Consider the β-linear divisor method and suppose that $\beta = k/2$ for some integer k. Then* (3.5) *holds for any rational $p_1, \dots, p_m$.*

Proof. By (3.14), the result holds for β if and only if it holds for $\beta + 1$. Hence it suffices to consider $\beta = 1/2$ and $\beta = 1$. The case $\beta = 1/2$ is part of Theorem 11.2.

For $\beta = 1$, we argue similarly. (Cf. the proof of Theorem A.11.) Let $0 \leq N < L$ and let $t \geq 0$ be such that $s_i(N) = [p_i t]_1$, see (2.7) and (A.3). Recall that $[p_i t]_1 = \lfloor p_i t \rfloor$ except that $\lfloor p_i t \rfloor - 1$ also is possible when $p_i t$ is an integer.

Since $p_i L$ is an integer, if we first for simplicity assume that $p_i t$ is not,

$$[p_i(2L - t)]_1 = 2p_i L - [p_i t]_1 - 1 = 2p_i L - s_i(N) - 1, \tag{11.2}$$

where summing over i yields

$$\sum_{i=1}^{m} [p_i(2L - t)]_1 = 2L - \sum_{i=1}^{m} s_i(N) - m = 2L - N - m. \tag{11.3}$$

Consequently, (A.3) again shows that these numbers yield the seat distribution for $2L - N - m$ seats. (Note that $L = \sum_{i=1}^{m} p_i L \geq m$, and thus $2L - N - m > 0$.) Hence, using (11.2) again,

$$s_i(2L - m - N) = [p_i(2L - t)]_1 = 2p_i L - s_i(N) - 1. \tag{11.4}$$

The argument works also, with a little care, if some $p_i t$ is an integer, and shows that (11.4) holds generally, and that ties in the distribution of $s_i(N)$ correspond to ties in the distribution of $s_i(2L - m - N)$. Hence, we have, regarding Δ_i as a random variable when there is a tie,

$$\Delta_i(2L-m-N) = s_i(2L-m-N) - (2L-m-N)p_i = mp_i - \Delta_i(N) - 1. \tag{11.5}$$

Taking the expectations in case of a tie, and then the average over the period $N = 0, \dots, L-1$, when $2L - m - N$ runs through the period $L - m + 1, \dots, 2L - m$, we see that if the average is x, then $x = mp_i - x - 1$; hence $x = \frac{1}{2}(mp_i - 1)$, and thus (3.5) holds when $\beta = 1$. ☐

For the 1-quota method (Droop), we have seen that (3.16) holds in Example 11.1; in fact, when $m = 2$, it holds for any (p_1, p_2) by Theorem 11.3, because the 1-quota method equals the 1-linear divisor method when $m = 2$, see Appendix B.1. However, it does not hold for $m = 3$.

Example 11.4. Let $m = 3$ and $(p_1, p_2, p_3) = (\frac{2}{5}, \frac{2}{5}, \frac{1}{5})$. Consider the 1-quota method (Droop). For $N = 1, 2, 3, 4, 5$, the smallest party gets, taking the average when there is a tie, $0, 0, 1, \frac{2}{3}, 1$ seats. (Note the non-monoticity, an instance of the Alabama paradox which frequently occurs for all quota methods, cf. [17].) The seat excesses are $-\frac{1}{5}, -\frac{2}{5}, +\frac{2}{5}, -\frac{2}{15}, 0$, with average $-\frac{1}{15}$, while (3.16) would give $p_1 - \frac{1}{3} = -\frac{2}{15}$. Thus (3.16) does not hold.

Note that this example shows that the heuristic argument in Remark 3.12 is not always valid when $p_1, \dots, p_m$ are rational. In this example, the retracted seat in Remark 3.12 is *not* taken uniformly; a calculation shows that it is taken from the smallest party only $\frac{4}{15}$ of the times.

We can ask the following for any β-linear divisor method or γ-quota method, where the cases $\gamma = 1$ in Problem 11.5 and $\beta = 1/2$, $\beta = 1$ and $\gamma = 0$ in Problem 11.6 are especially interesting. (By Theorem A.12, for the linear divisor methods, it suffices to consider $0 < \beta \leq 1$.)

Problem 11.5. Find a general formula for the asymptotic bias $\lim \mathbb{E}\,\Delta_i$ for rational $p_1, \ldots, p_m$. (Theorems 11.2 and 11.3 yield the answers for $\beta \in \mathbb{Z}/2$ and $\gamma = 0$.)

Problem 11.6. Find a general formula for the asymptotic distribution of Δ_i for rational $p_1, \ldots, p_m$.

Appendix A. Divisor methods

A *divisor method* is based on a given sequence of numbers $d(n)$, $n \geq 1$, with $0 \leq d(1) < d(2) < d(3) < \ldots$. Different choices of $d(n)$ give different methods. A number of divisor methods that have been used or discussed are shown in Table 1. The most important ones are the widely used methods by Jefferson/D'Hondt and Webster/Sainte-Laguë. (Notation varies. In e.g. [1], the sequence is denoted $d(0), d(1), \ldots$; thus their $d(n)$ is our $d(n+1)$.)

Divisor methods can be described in several different ways that yield the same result. The methods have been invented and reinvented in different guises; tradition varies, and different types of formulations are used in, for example, election laws and other official and non-official descriptions from different countries. There are two main types of formulation, and we give each in two different versions. (The equivalence of the different types of formulations have been known for a long time. For example, D'Hondt proposed his method in [5] using essentially our formulation D1 below, and later showed the equivalence to D4, see [18, p. 125].)

In the first type of formulation, the method is seen as a way of rounding. Given a sequence $d(1), d(2), \ldots$ as above, we define the *d-rounding* of a real number $x > 0$ by

$$[x]_d := n \qquad \text{if} \quad d(n) \leq x \leq d(n+1), \tag{A.1}$$

where $d(0) = 0$. Note that this is unambiguous only if $d(n) < x < d(n+1)$; if $x = d(n)$ for an integer n, then both $[x]_d = n$ and $[x]_d = n-1$ are acceptable values. (This is important in the case of ties, see below.) In particular, $d(n) = n - \frac{1}{2}$ (Webster's method) yields standard rounding; $d(n) = n$ (Jefferson's method) yields rounding downwards; $d(n) = n-1$ (Adams's method) yields rounding upwards. (Other choices of $d(n)$ yield non-standard roundings. Note that "rounding" in general should be interpreted in a very weak sense, and that $|\,[x]_d - x| > 1$ may be possible, see Remark A.3.) In this formulation, the numbers $d(n)$ are sometimes called *signposts*, see [1] and [8].

Using the concept of d-rounding, and the notation from Section 2, the divisor method may be defined as follows:

	$d(n)$	decimal values (rounded)
Jefferson, D'Hondt	n	1, 2, 3, 4, ...
Webster, Sainte-Laguë	$n - \frac{1}{2}$	0.5, 1.5, 2.5, 3.5, ...
Adams	$n - 1$	0, 1, 2, 3, ...
Imperiali	$n + 1$	2, 3, 4, 5, ...
Danish	$n - \frac{2}{3}$	0.333, 1.333, 2.333, 3.333, ...
Adjusted Sainte-Laguë	$\begin{cases} 0.7, & n = 1 \\ n - \frac{1}{2}, & n \geq 1 \end{cases}$	0.7, 1.5, 2.5, 3.5, ...
Cambridge Compromise	$(n-6)_+$	0, 0, 0, 0, 0, 0, 1, 2, 3, ...
Huntington	$\sqrt{n(n-1)}$	0, 1.414, 2.449, 3.464, ...
Dean	$\frac{2n(n-1)}{2n-1}$	0, 1.333, 2.400, 3.429, ...
Estonia	$n^{0.9}$	1, 1.866, 2.688, 3.482, ...
Macau	2^{n-1}	1, 2, 4, 8, ...

TABLE 1. Some divisor methods. (For Dean's method, $d(n)$ is the harmonic mean of n and $n-1$; cf. Webster's and Huntington's methods with the arithmetic and geometric means.)

Divisor method, formulation D1. Let

$$s_i := \left[\frac{v_i}{D}\right]_d, \tag{A.2}$$

where D is a (real) number chosen such that $\sum_{i=1}^m s_i = N$.

The rationale for this formulation is that D is regarded as the price of a seat, i.e. the number of votes that a seat "costs", which would give v_i/D seats to party i; this real number is d-rounded to an integer to obtain s_i. The price D is set by "the market" (i.e. the election officer, or computer) so that the desired total number of seats is distributed. This interpretation can be further combined with the definition (A.1) to say that the price of n seats is $d(n)D$ votes.

Remark A.1. In particular, (A.1) yields $[x]_d = 0$ if and only if $x < d(1)$, or perhaps $x = d(1)$. Thus, if $d(1) = 0$, then $[x]_d \geq 1$ for every $x > 0$, while if $d(1) > 0$, then $[x]_d = 0$ for small positive x.

Hence, if $d(1) = 0$, then (A.2) yields $s_i \geq 1$, so every party (with at least one vote) is guaranteed at least one seat. This is unacceptable in general elections, where on the contrary there often are special threshold rules to prevent small parties from getting seats. However, this is acceptable, and may even be desirable, when distributing seats between states or constituencies, where typically there is a side condition of at least one seat each (and sometimes a minimum of two or more seats); a typical example is the United States Constitution [1]. (If $d(1) = 0$, we assume that $N \geq m$, so that there are enough seats to satisfy this minimum requirement.)

A divisor method is by [8] called *impervious* if $d(1) = 0$ and *pervious* if $d(1) > 0$.

If we move D continuously from ∞ down to 0, it is seen that the total number of seats $\sum_{i=1}^m s_i$ given by (A.2) grows monotonously from 0 (when $d(1) > 0$) or m (when $d(1) = 0$) to ∞, and there is a value of D that yields the desired sum $\sum_{i=1}^m s_i = N$. In general there is an interval of such D, all yielding the same seat distribution s_i, and this determines all s_i uniquely. In exceptional cases, there is only a single D that works; in this case we must have equalities $v_i/D = d(s_i)$ for some party i and $v_j/D = d(s_j - 1)$ for some other party j, see (A.2) and (A.1), and in this case the result $(s_i)_1^m$ is not unique; we may move one seat from party j to party i and (A.2) still holds. Such cases are known as *ties*; they are usually resolved by drawing lots, but other special rules for ties are used in some countries. The formulation D1 thus gives a well-defined method in the sense that the numbers s_i are uniquely determined, except possibly when there are ties. (Note that the number D is not uniquely determined.)

It is sometimes, as in Section 4, convenient to rewrite (A.2) as follows using $x := 1/D$ or $t := V/D$. A divisor method may thus also be called a *multiplier method*, as in [13].

Divisor method, formulation D2. Let

$$s_i := [v_i x]_d = [p_i t]_d\,, \tag{A.3}$$

where x or t is a (real) number chosen such that $\sum_{i=1}^m s_i = N$.

Formulations of the type D1 of divisor methods are the standard in USA, where Thomas Jefferson formulated a method of this type (with rounding downwards) in 1792 for apportionment; Jefferson's method was adopted by Congress in 1792 and used until 1832 [1]. Four other divisor methods (using formulation D1 with different $d(n)$, see Table 1) have also been important in USA; Huntington's method is used since 1941 and Webster's method has been used earlier, while Adams's and Dean's methods have never been used but have frequently figured in discussions, see [15] and [1].

In Europe, this type of formulation is unusual (although [10] and the current German election law [35, § 6 (2)] are examples), and divisor methods are more commonly defined by formulations such as the two following (closely related) ones. Note that these formulations are algorithmic, in the sense that they directly yield workable methods, while with D1 one has to search for a suitable divisor. (Preferably by computer, where such searches are easy.) The European tradition of divisor methods is younger than the American one, see e.g. [4], [7], [18]. D'Hondt's method (which is the same as Jefferson's) was proposed by D'Hondt [5] in 1882 and first adopted in Belgium 1899; it is now used in many countries. Sainte-Laguë's method (which is the same as Webster's) was proposed by Sainte-Laguë [26, 27] in 1910, and is also used in several places.

Divisor method, formulation D3. The N seats are awarded sequentially, with each seat given to the party i that has the highest value of the quotient $v_i/d(s_i+1)$ where s_i is the number of seats that the party has received so far.

The quotient $v_i/d(s_i+1)$ is called the *comparative figure* of the party. Note that the comparative figure is updated (decreased) each time the party receives a seat.

Divisor method, formulation D4. Divide each v_i by the numbers $d(1)$, $d(2)$, ... (as far as necessary). Assign the seats to the N largest quotients $v_i/d(j)$.

With formulation D4, the quotients $v_i/d(j)$ can be arranged in a table (matrix); alternatively, the sequences of quotients obtained for different parties can be merged into a single sequence in decreasing order. It is immediate that formulations D3 and D4 are equivalent; D3 picks the largest quotients in D4 in decreasing order.

Of course, ties can occur in formulations D3 and D4 as well, and again they are resolved by drawing lots, or possibly by some other rule. (In formulation D3, ties may also occur at intermediate stages, but they do not affect the final result.)

To see that the formulations D1–D4 are equivalent and yield the same method, note first that (A.2) can be written, using the definition (A.1),

$$d(s_i) \le \frac{v_i}{D} \le d(s_i+1), \qquad i=1,\dots,N, \tag{A.4}$$

or, equivalently (interpreting $v_i/0=+\infty$)

$$\frac{v_i}{d(s_i+1)} \le D \le \frac{v_i}{d(s_i)}, \qquad i=1,\dots,N. \tag{A.5}$$

Consequently, $(s_i)_{i=1}^N$ are such that

$$\max_i \frac{v_i}{d(s_i+1)} \le \min_i \frac{v_i}{d(s_i)}; \tag{A.6}$$

conversely, for any such s_i we may choose D such that (A.5) holds. Thus formulation D1 is equivalent to:

Divisor method, formulation D5. $s_1,\dots,s_m$ are chosen such that $\sum_{i=1}^m s_i = N$ and (A.6) holds.

Furthermore, it is easily seen that formulation D5 is equivalent to D4, and thus also to D3. Hence all formulations are equivalent. (It is easily verified that also ties appear simultaneously in the different formulations, except that in formulation D3 there may be additional intermediate ties that do not affect the result.) Further equivalent formulations are given by [8].

We see also, from (A.5), that the possible choices of D in formulation D1 that yield $s_1,\dots,s_m$ are $D\in[D_-,D_+]$, where

$$D_- := \max_i \frac{v_i}{d(s_i+1)}, \qquad D_+ := \min_i \frac{v_i}{d(s_i)}. \tag{A.7}$$

Thus, D_+ equals the winning comparative figure when the last seat is awarded in formulation D3, and D_- equals the winning comparative figure for the next seat (if there were one).

Remark A.2. It follows immediately from any of the formulations D1–D5 that the sequences $d(n)$ and $cd(n)$, for any constant $c > 0$, define the same divisor method. For example, Sainte-Laguë's method is given in Table 1 with $d(n) = n - \frac{1}{2}$; it can as well be defined with $d(n) = 2n - 1$. (In fact, the method is traditionally defined with this sequence, see e.g. Sainte-Laguë [26, 27]. It is therefore also called the *odd-number method*.) Similarly, the adjusted Sainte-Laguë's method (used in Sweden and Norway) is traditionally defined with the sequence 1.4, 3, 5, 7, . . . , and the Danish method with the sequence 1, 4, 7, 10, ... (This method is used in Danish parliamentary elections for part of the distribution of adjustment seats between consituencies; the main distribution among parties is by the method of largest remainder.) It may be convenient to use integers; on the other hand, Imperiali's method (used in Belgian local elections) is in the election law [33, Art. 56] described using 1, $1\frac{1}{2}$, 2, $2\frac{1}{2}$, . . . , i.e., $(n + 1)/2$. In Table 1, we have normalised $d(n)$ to the form $n + \beta$ whenever possible, cf. Section A.1.

Remark A.3. Several papers impose the condition that $n - 1 \le d(n) \le n$, but this is really not necessary. This condition is natural for the interpretation of the method as a kind of rounding in formulation D1, since it is equivalent to $[n]_d = n$ for any integer $n \in \mathbb{N}$, see (A.1). In particular, this means that if the exact proportions $q_i = p_i N$ happen to be integers, then the divisor method gives $s_i = q_i$. (Take $D = V/N$ in (A.2); cf. (2.1).) However, the method is well-defined also for other sequences $d(n)$, and as seen in Table 1, there are some divisor methods currently in use for general elections that do not satisfy this condition (the Imperiali method [33, Art. 56] and the methods in Estonia [34, § 62 (5)] and Macau [36, Artigo 17]). Formulation D1 remains formally valid in these cases too, but in practice these methods are described using e.g. formulation D3 where there are no problems of interpretation for any increasing sequence $d(n)$. (Methods satisfying $s_i = q_i$ when all q_i are integers are called *weakly proportional* by [1].)

Remark A.4. We required above that the sequence $d(n)$ is strictly increasing. In fact, it can be defined assuming only $0 \le d(1) \le d(2) \le \dots$, with only a few minor modifications above. This extension is hardly used in practice, except that a minimum requirement of at least r seats for each party for some $r > 1$ can be achieved by taking $d(1) = d(2) = \dots = d(r) = 0$, cf. Remark A.1 for the case $r = 1$. (We then need $N \ge rm$.)

For example, the Cambridge Compromise [10] proposed for apportionment of the European Parliament can be described as giving each state a "base" of 5 seats and distributing the rest by Adams's method (subject to a maximum of 96 seats for each state); this is (e.g. by formulation D1, see also Theorem A.12 below) equivalent to giving each state 6 seats and distributing the rest by D'Hondt's method (subject to the same maximum). If we ignore the restriction to at most 96 seats (as done in Table 1 for simplicity), this is the same as a divisor method with $d(n) = 0$ for $n \le 6$ and $d(n) = n - 6$ for $n > 6$, i.e. $d(n) = (n - 6)_+$. The maximum restriction is easily implemented if e.g. formulation D3 is used, cf. [10]; formally we can also implement it by redefining $d(n) = \infty$ for $n > 96$.

Remark A.5. The name "divisor method" is used for methods of this type in any of the formulations above. However, the word *divisor* is usually used for the (variable) number D in formulation D1, but for the (fixed) numbers $d(1), d(2), \dots$ in formulations D3–D4.

A.1. **Linear divisor methods.** We say that a divisor method is *linear* if $d(n) = an + b$ for some $a > 0$ and $b \in \mathbb{R}$. By Remark A.2, we may replace $d(n)$ by $d(n)/a = n + b/a$. We define $\beta := b/a + 1$. In other words, we may assume that $a = 1$ and

$$d(n) = n - 1 + \beta \tag{A.8}$$

for some real β. The reason for our choice of β as parameter, and thus formula (A.8), is (A.9) below. Note that $\beta = d(1)$.

We call the divisor method with $d(n)$ given by (A.8) the *β-linear divisor method*. The parameter β is called *proportionality index* (*Proportionalitätsindex*) by [18]. The method is called *q-stationary multiplier method* (where $q = \beta$) by [13].

With a linear divisor method (A.8), the definition (A.1) of d-rounding yields that if $[x]_d = n$, then $n - 1 + \beta \le x \le n + \beta$, or equivalently $x - \beta \le n \le x - \beta + 1$. Consequently, d-rounding equals the β-rounding defined in (2.4) and (2.5); i.e.

$$[x]_d = [x]_\beta\,, \tag{A.9}$$

at least provided $x > 0$ and $x > \beta - 1$. Hence, the β-linear divisor method can, using formulation D1, be defined by (2.7) as we did in Section 2. If $\beta > 1$, we here have to assume that N is so large that $v_i/D > \beta - 1$ and thus $[v_i/D]_\beta \ge 0$.

Remark A.6. If $\beta > 1$ and N is small, (2.7) may give a negative seat number s_i for a very small party. Of course, this has to be replaced by 0, as given by (A.2). Hence, we may still use (2.7) if we ignore any party that would get a negative number of seats.

If $[v_i/D]_\beta \le -1$ for some i, then $v_i/D \le \beta - 1$ by (2.5) and thus for any party j, using (2.4),

$$\Bigl[\frac{v_j}{D}\Bigr]_\beta \le \Bigl[\frac{v_j}{v_i}(\beta - 1)\Bigr]_\beta = \Bigl[\frac{p_j}{p_i}(\beta - 1)\Bigr]_\beta \le \Bigl(\frac{p_j}{p_i} - 1\Bigr)(\beta - 1),$$

and summing over j we get $N \le (1/p_i - m)(\beta - 1)$. Hence, if $N > (1/\min_i p_i - m)(\beta - 1)$, then (2.7) holds without exception.

Remark A.7. The definition above requires $\beta = d(1) \ge 0$. We can extend it to $\beta < 0$ by replacing (A.8) by

$$d(n) = (n - 1 + \beta)_+, \tag{A.10}$$

i.e., letting $d(n) = 0$ for $n \le 1 + \lfloor |\beta| \rfloor$, cf. Remark A.4. Then (A.9) still holds for $x > 0$, and (2.7) holds as before. By Remark A.4, every party now gets at least $\lfloor |\beta| \rfloor + 1$ seats, and thus we need $N \ge m\bigl(\lfloor |\beta| + 1 \rfloor\bigr)$.

The Cambridge Compromise (without maximum rule) is an example, with $\beta = -5$.

Remark A.8. Linear divisor methods are often considered only for the case $0 \le \beta \le 1$, which is equivalent to the condition $n - 1 \le d(n) \le n$ discussed in Remark A.3. However, we see that they are well-defined for arbitrary real β, although some care may be needed for small N. Note that at least the case $\beta = 2$ is used in practice.

In the main part of the paper we consider only linear divisor methods. Note that many of the methods in Table 1 are linear: Jefferson/D'Hondt ($\beta = 1$); Webster/Sainte-Laguë ($\beta = 1/2$); Adams ($\beta = 0$); Imperiali ($\beta = 2$); Danish ($\beta = 1/3$); Cambridge Compromise ($\beta = -5$).

Remark A.9. The adjusted Sainte-Laguë method also has a linear $d(n)$ as in (A.8) ($\beta = 1/2$) except for $d(1)$, which does not affect asymptotic results. More precisely, the adjusted Sainte-Laguë method differs from Sainte-Laguë's method only in the value of $d(1)$, and $[x]_d$ is the same for both methods for all $x \ge 0.7$; hence they give exactly the same result as soon as every party gets at least one seat by the adjusted method. (But the adjustment makes it more difficult for a small party to get the first seat.) For asymptotic results as in the present paper, there is thus no difference between the adjusted Sainte-Laguë method and Sainte-Laguë's method.

Remark A.10. Huntington's and Dean's methods are asymptotically linear in the sense that $d(n) = n-1+\beta+o(1)$ as $n \to \infty$, in both cases with $\beta = 1/2$. An argument similar to the proof in Section 4 shows that as $N \xrightarrow{\text{p*}} \infty$, the probability that these methods yield the same result as Webster's tends to 1. In particular, Theorem 3.7 (with $\beta = 1/2$) holds for these methods too, and Theorem 3.4 shows that they are asymptotically unbiased.

Note that our asymptotic approach does not answer the controversial question whether Webster's or Huntington's method is the most fair and unbiased, see [1]; this depends mainly on how one measures the bias for small parties (states), in particular for the ones obtaining just 1 seat.

Although linear divisor methods may be biased, see Section 3, they are perfectly proportional with respect to changes in the total number of seats. (If $\beta < 0$ or $\beta > 1$, we assume that N is not too small, see Remarks A.6–A.7.)

Theorem A.11. *Consider a β-linear divisor method. If the number of seats is increased from N to $N + L$, and $Lp_1, \dots, Lp_m$ all are integers, then party i gets exactly Lp_i seats more:*

$$s_i(N + L) = s_i(N) + Lp_i. \tag{A.11}$$

Proof. This is easiest seen using the multiplier formulation D2. If t yields $s_i(N) = [p_i t]_\beta$ with $\sum_{i=1}^m s_i(N) = N$, then $t + L$ yields

$$s_i = [p_i(t + L)]_\beta = [p_i t]_\beta + p_i L = s_i(N) + p_i L,$$

with $\sum_{i=1}^m s_i = \sum_{i=1}^m s_i(N) + \sum_{i=1}^m p_i L = N + L$. □

There is also a simple relation between the methods for two different values of β that differ by an integer.

Theorem A.12. *Consider the β-linear divisor method and suppose that N is so large that every party gets at least one seat. (Otherwise, ignore the remaining parties.) Then the β-linear divisor method yields the same result as first giving one seat to each party and then distributiong the remaining $N - m$ seats by the $(\beta + 1)$-linear divisor method.*

Proof. This follows immediately from formulation D4, noting that by (A.8), or more generally (A.10), $d_{\beta+1}(n) = d_\beta(n + 1)$ (using β as a subscript to denote the divisor sequences).

Alternatively, this follows from (2.7) and the relation $[x]_{\beta+1} = [x]_\beta - 1$, which follows from (2.6). □

By repeated applications of this theorem, we can reduce any β-linear divisor method to the case $0 < \beta \leq 1$.

Example A.13. Adams's method ($\beta = 0$) is the same as giving each party one seat and distributing the rest by Jefferson's method ($\beta = 1$).

Example A.14. The Imperiali method ($\beta = 2$) is equivalent to using D'Hondt's method ($\beta = 1$) but retracting one seat from each party and redistribute them (among the parties that had at least one seat) by continued application of D'Hondt's method.

Example A.15. The Cambridge Compromise ($\beta = -5$), is as said in Remark A.4 the same as giving each party (state) 5 seats and distributing the rest by Adams's method ($\beta = 0$), or as giving each party (state) 6 seats and distributing the rest by Jefferson/D'Hondt's method ($\beta = 1$).

A.2. **Divisor methods and proportionality.** We make a formal definition of proportionality for election methods, using an asymptotic property, since no election method is exactly proportional except in exceptional cases. (Balinski and Young [1] uses a different definition, cf. Remark A.3.)

Definition A.16. An election method is *(asymptotically) proportional* if, for any number m of parties and any proportions of votes $p_1, \dots, p_m$, for each party i,

$$s_i(N)/N \to p_i \qquad \text{as } N \to \infty. \tag{A.12}$$

Many proportional methods that are used are divisor methods, but not all divisor methods are proportional. In order to characterise the proportional divisor methods, recall that a positive measurable function f on some interval (A,∞) is *regularly varying with index* ρ, where ρ is a real number, if

$$f(\lambda x)/f(x) \to \lambda^\rho \qquad \text{as } x\to\infty, \tag{A.13}$$

for every $\lambda>0$. In the special case $\rho=0$, i.e. when $f(\lambda x)/f(x)\to 1$ as $x\to\infty$ for every $\lambda>0$, we say that f is *slowly varying.* A typical example of a slowly varying function is $\log x$. A sequence c_n is said to be regularly varying with index ρ (and slowly varying if $\rho=0$) if the function $c_{\lfloor x\rfloor}$ is; this is equivalent to $c_{\lfloor \lambda n\rfloor}/c_n\to\lambda^\rho$ as $n\to\infty$, for every $\lambda>0$. (See [3], where many more results are given.)

Theorem A.17. *For a divisor method defined by $d(1),d(2),\dots$, the following are equivalent.*

(i) *The method is proportional.*
(ii) *For any m and $p_1,\dots,p_m$, and all $i,j\le m$,*

$$\frac{s_j(N)}{s_i(N)}\to\frac{p_j}{p_i} \qquad \textit{as } N\to\infty. \tag{A.14}$$

(iii) *$x\mapsto [x]_d$ is regularly varying with index 1.*
(iv) *$x\mapsto [x]_d/x$ is slowly varying.*
(v) *The sequence $d(n)$ is regularly varying with index 1.*
(vi) *The sequence $d(n)/n$ is slowly varying.*

The proof shows that it suffices that (A.12), or (ii), holds when $m=2$.

Proof. (i) $\iff$ (ii). If (A.12) holds for each i, then also (A.14) holds. On the other hand, if (A.14) holds, then summing (A.14) over all j we obtain $N/s_i(N)\to 1/p_i$, and thus (A.12).

(ii) $\iff$ (iii). By (A.2), $s_i=[v_i/D]_d=[p_ix]_d$ with $x:=V/D$, and hence (A.14) can be written

$$\frac{[p_jx]_d}{[p_ix]_d}\to\frac{p_j}{p_i} \qquad \text{as } x\to\infty, \tag{A.15}$$

since $N=\sum_{i=1}^m s_i=\sum_{i=1}^m [p_ix]_d$ and thus $N\to\infty \iff x\to\infty$. If $[x]_d$ is regularly varying with index 1, then (A.15) follows from (A.13) with $\lambda=p_j/p_i$. Conversely, if (ii) holds, and $\lambda>0$, consider the case $m=2$ and $p_1=\lambda/(1+\lambda)$, $p_2=1/(1+\lambda)$. Then (A.15) with $(j,i)=(1,2)$ and $y:=p_2x$ yields $[\lambda y]_d/[y]_d\to\lambda$, which is (A.13) with $\rho=1$.

(iii) $\iff$ (iv). Directly by the definition (A.13).

(iii) $\iff$ (v). Let $g(x)$ be the inverse of $[x]_d$ defined by $g(x):=\inf\{y: [y]_d>x\}$. Then (A.1) shows that $g(x)=d(\lfloor x\rfloor+1)$. [3, Theorem 1.5.12] shows that if $[x]_d$ is regularly varying with index 1, then so is $g(x)$. This implies $d(n+1)/d(n)\to 1$ as $n\to\infty$, and it follows that $d(n)$ is regularly varying with index 1. The converse follows in the same way, now using the inverse of $d(\lfloor x\rfloor)$ which is $[x]_d+1$.

(v) $\iff$ (vi). Directly by the definition. ∎

In particular, Theorem A.17 shows that any divisor method with

$$d(n)/n \to a \qquad \text{as } n \to \infty, \tag{A.16}$$

for some $a > 0$, is proportional. Of the methods in Table 1, the only methods that do not satisfy (A.16), and thus are proportional, are the methods of Estonia and Macau. In fact, it is easily seen from formulation D3 or D4 that the Estonian method is the same as D'Hondt's method applied to $v_i^{1/0.9} = v_i^{1.111\cdots}$; thus it makes s_i proportional to $v_i^{1.111\cdots}$. This gives a small but clear advantage to larger parties. Macau's method, on the contrary, favours small parties and encourages splintering. (With only 12 directly elected members of the Legislative Assembly of Macau, it essentially imposes a maximum of 2 seats per party.)

In fact, the proportional divisor methods in Table 1, and all linear divisor methods, satisfy the following stronger form of proportionality.

Theorem A.18. *For a divisor method with $d(n) = an + O(1)$ for some $a > 0$, $s_i(N) = p_i N + O(1)$. (The implicit constant may depend on m, but not on anything else.)*

Proof. By Remark A.2, we may replace $d(n)$ by $d(n)/a$, and thus assume $d(n) = n + O(1)$.

If $d(n) = n + O(1)$, then also $d(n-1) = n + O(1)$ and thus (A.1) shows that if $[x]_d = n$, then $x = n + O(1)$; in other words,

$$[x]_d = x + O(1). \tag{A.17}$$

Consequently, (A.2) yields, with $y := V/D$,

$$s_i = \frac{v_i}{D} + O(1) = p_i y + O(1). \tag{A.18}$$

Summing over i we find $N = \sum_{i=1}^m p_i y + O(1) = y + O(1)$, and thus (A.18) yields $s_i = p_i N + O(1)$. □

A.3. **Uniformity.** Divisor methods satisfy the following consistency property, called *uniformity* in [1], when we consider subsets of the set of all parties. This is obvious from any of the formulations D1–D5, and would perhaps not be worth mentioning, except for the fact that it does *not* hold for quota methods; in fact, under some weak extra conditions, it is satisfied only by divisor methods [1].

Theorem A.19. *Suppose that some set of parties $A_1, \dots, A_\ell$ (where $1 \le \ell \le m$) get together N' seats in an election by a divisor method. Then the number of seats for each of these parties is the same as if N' seats were distributed, by the same divisor method, in an election with only these parties participating (and obtaining the same numbers of votes v_i).* □

Appendix B. Quota methods

In a *quota method* one first determines a *quota* Q; different quota methods differ in the formula for Q, see below. (Q can be seen as the standard price of a seat, just as the divisor D in the formulation D1 of divisor methods.) The seats are then distributed as follows:

Quota method, formulation Q1. Divide the numbers of votes v_i by Q, and give first each party as many seats as the integer part $\lfloor v_i/Q \rfloor$ of its fraction. Any remaining seats are given to the parties that have largest fractional part $\{v_i/Q\} = v_i/Q - \lfloor v_i/Q \rfloor$. (Or, equivalently, largest remainder at the division v_i/Q, since the remainder is $v_i - \lfloor v_i/Q \rfloor Q$, which is Q times the fractional part $v_i/Q - \lfloor v_i/Q \rfloor$.)

The most common quota method is the *method of largest remainder*, or *Hamilton's method*, in Europe often called *Hare's method*, which uses the *simple quota*, also called *Hare quota*, $Q := V/N$, i.e. the average number of votes per seat. Note that then $v_i/Q = q_i$, the exactly proportional real allocation, see (2.1). (Alexander Hamilton suggested the method in 1792 for apportionment to the US Congress; it was approved by Congress but vetoed by president Washington; it made a comeback and was used (under the name *Vinton's method*) 1850–1900, see [1]. The namn Hare's method is also well-established but is a misnomer; Thomas Hare really advocated a different method, the *Single Transferable Vote, STV*, which is not based on party lists and is not treated here.)

Another version is *Droop's method* which uses the *Droop quota* $Q := V/(N+1)$. Also the *Imperiali quota* $Q := V/(N+2)$ has occasionally been used (e.g. in Italy until 1993 [7]).

We define, more generally, for any real number γ, the *γ-quota method* to be the quota method with quota

$$Q := \frac{V}{N+\gamma}. \tag{B.1}$$

(For $\gamma < 0$, we assume that $N > |\gamma|$.) Thus Hamilton/Hare's method is the case $\gamma = 0$, Droop's method is $\gamma = 1$, and the Imperiali quota is $\gamma = 2$. (Kopfermann [18] calls the γ-quota method *Rundungsverfahren mit dem Proportionalitätsindex* ρ, where his $\rho = (\gamma+1)/2$, i.e., $\gamma = 2\rho - 1$.)

Remark B.1. The quotas above are sometimes rounded to integers (upwards, downwards or by standard rounding); see [22] for several different examples in current use. There is no mathematical reason to round the quota; on the contrary, there are reasons against. Kopfermann [18] points out that quota methods involve rounding at the end, but that, as a general principle, intermediate values in a calculation should not be rounded. A specific problem is that rounding the quota means that the method is no longer homogeneous (see Section 2). Moreover, with a rounded quota, in some cases a party may lose a seat by getting an additional vote, which ought to be unacceptable; see [16] for a simple example with Droop's method (although there formulated for STV, which in the example gives the same result). In a general election with a large number of votes, and therefore a large quota, rounding the quota will usually make no difference, but in the exceptional cases when it does make a difference, it can be harmful and lead to undesirable results.

In the present paper, we consider the methods with unrounded quotas.

Remark B.2. In formulation Q1, it is implicitly assumed that $N - m \leq \sum_{i=1}^{m} \lfloor v_i/Q \rfloor \leq N$, so that the number of remaining seats is non-negative and not larger than m, the number of parties. For the γ-quota method, this is always satisfied if $-1 \leq \gamma < 1$ (including Hamilton/Hare), and also if $\gamma = 1$ (Droop) except in a very special case when all parties tie for the last seat. For $\gamma < -1$ or $\gamma > 1$, formulation Q1 may have to be amended; in such cases we interpret it by using one of the equivalent formulations Q2–Q4 below. (These formulations are well-defined for all Q and N, up to the usual possibility of ties.)

The rule in Q1 that the remaining seats are distributed to the parties with largest remainders $v_i/Q - \lfloor v_i/Q \rfloor$ can also be expressed by saying that we find a suitable α and round v_i/Q upwards if the fractional part is greater than α, and downwards if the fractional part is smaller. In other words, recalling the notion of α-rounding in Section 2, a quota method can equivalently be described as follows.

Quota method, formulation Q2.

$$s_i := \left[\frac{v_i}{Q}\right]_\alpha, \tag{B.2}$$

where α is chosen such that $\sum_{i=1}^{m} s_i = N$.

For the γ-quota method we have $Q = V/(N + \gamma)$, and thus $v_i/Q = p_i V/Q = p_i(N + \gamma)$. Hence we have also the following formula.

Quota method, formulation Q3. The γ-quota method is given by

$$s_i := [(N + \gamma)p_i]_\alpha\,, \tag{B.3}$$

where α is chosen such that $\sum_{i=1}^{m} s_i = N$.

Formulations Q2–Q3 are the same as (2.8).

Quota methods can also be described using a sequential allocation of seats by comparative figures as in formulation D3 of divisor methods. (This is not the usual way to define quota methods, but it appears in proofs in [27] and [23].) It is easy to see that Q2 is equivalent to the following.

Quota method, formulation Q4. The N seats are awarded sequentially, with each seat given to the party i that has the highest value of the difference $v_i - s_i Q$ where s_i is the number of seats that the party has received so far.

It is easy to see that every γ-quota method satisfies $s_i(N) = Np_i + O(1)$, uniformly in all m and $p_1, \dots, p_m$, and hence is proportional, cf. Definition A.16 and Theorem A.18. We omit the easy proof since we show a precise version of this in Theorem 3.10.

Moreover, a γ-quota method is perfectly proportional with respect to changes in the total number of seats, just as a β-linear divisor method, cf. Theorem A.11.

Theorem B.3. *Consider a γ-quota method. If the number of seats is increased from N to $N + L$, and $Lp_1, \dots, Lp_m$ all are integers, then party i gets exactly Lp_i seats more:*

$$s_i(N + L) = s_i(N) + Lp_i. \tag{B.4}$$

Proof. Obvious from (B.3), since the same α works for N and $N + L$. □

B.1. **Two parties.** In the simple case $m = 2$, when there are only two parties, the β-linear divisor method and the γ-quota method with $\gamma = 2\beta - 1$ coincide [18, Satz 6.2.6]. (In particular, then Webster/Sainte-Laguë = Hamilton/Hare and Jefferson/D'Hondt = Droop.) In fact, it is easy to see that

$$s_i := [(N + 2\beta - 1)p_i]_\beta, \qquad i = 1, 2, \tag{B.5}$$

gives $s_1 + s_2 = N$, (at least for some choice in case of a tie). Thus, for the β-linear divisor method we can take $D' = 1/(N + 2\beta - 1) = 1/(N + \gamma)$ and $D = D'V = V/(N + 2\beta - 1)$ in (2.7), and for the γ-quota method we can take $\alpha = \beta$ in (2.8). Hence, both methods yield the result given by (B.5).

Appendix C. Lemmas on uniform distribution

Proof of Lemma 5.1. First, by Lemma 4.1, the sequence

$$(\{ny_1/m + a_1/m\}, \dots, \{ny_k/m + a_k/m\}) \tag{C.1}$$

is uniformly distributed in $[0, 1)^k$. Thus, by the definition of Mod_m,

$$\big(\mathrm{Mod}_m(ny_1 + a_1), \dots, \mathrm{Mod}_m(ny_k + a_k)\big) \tag{C.2}$$

is uniformly distributed in $[0, m)^k$. If a sequence (x_n) is uniformly distributed in $[0, m)$, then the sequence of pairs $(\{x_n\}, \lfloor x_n \rfloor)$ is uniformly distributed in $[0, 1) \times \{0, \dots, m - 1\}$. This extends to k dimensions, and thus it follows from (C.2) that if $\tilde{\ell}_{nj} := \mathrm{Mod}_m(\lfloor ny_j + a_j \rfloor) = \lfloor \mathrm{Mod}_m(ny_j + a_j) \rfloor$, then

$$\big(\{ny_1 + a_1\}, \dots, \{ny_k + a_k\}, \tilde{\ell}_{n1}, \dots, \tilde{\ell}_{nk}\big) \tag{C.3}$$

is uniformly distributed in $[0, 1)^k \times \{0, \dots, m - 1\}^k$.

The argument just given shows that (C.3) is uniformly distributed also along any subsequence $n = m\nu + n_0$, $\nu \geq 1$. We have $\ell_n = \mathrm{Mod}_m(n - \sum_{j=1}^k \tilde{\ell}_{nj})$, and it follows that along any such subsequence, ℓ_n is determined by $\tilde{\ell}_{n1}, \dots, \tilde{\ell}_{nk}$. Consequently, the vectors (5.11) are uniformly distributed along any such subsequence, and therefore also along the full sequence. □

The following lemma is essentially taken from [12]. (There Riemann integrability is assumed, but as the proof below shows, Lebesgue integrability suffices.)

Lemma C.1. *Let $X = (X_1, \dots, X_k)$ be a random variable with an absolutely continuous distribution in $\mathbb{R}^k$. Let ν_n be any sequence of constants with $\nu_n \to \infty$. Then*

$$\bigl(\{\nu_n X_1\}, \dots, \{\nu_n X_k\}, X_1, \dots, X_k\bigr) \xrightarrow{\mathrm{d}} \bigl(U_1, \dots, U_k, X_1, \dots, X_k\bigr) \tag{C.4}$$

where $U_1, \dots, U_k \sim \mathrm{U}(0,1)$ are independent of each other and of X.

Proof. The assumption that X is absolutely continuous means that it has a density function f, which is a Lebesgue integrable function on $\mathbb{R}^k$.

We regard $\{\nu_n X_j\}$ and U_j as elements of the circle group $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Thus the random vectors in (C.4) are elements of the group $\mathbb{T}^k \times \mathbb{R}^k$, and by standard Fourier analysis, it suffices to show that for any integers $\ell_1, \dots, \ell_k$ and real numbers $t_1, \dots, t_k$,

$$\mathbb{E}\, e^{\sum_j 2\pi \mathrm{i} \ell_j \{\nu_n X_j\} + \sum_j \mathrm{i} t_j X_j} \to \mathbb{E}\, e^{\sum_j 2\pi \mathrm{i} \ell_j U_j + \sum_j \mathrm{i} t_j X_j}. \tag{C.5}$$

Since $\nu_n X_j - \{\nu_n X_j\}$ is an integer, we have, with f as above,

$$\begin{aligned}\mathbb{E}\, e^{\sum_j 2\pi \mathrm{i} \ell_j \{\nu_n X_j\} + \sum_j \mathrm{i} t_j X_j} &= \mathbb{E}\, e^{\sum_j 2\pi \mathrm{i} \ell_j \nu_n X_j + \sum_j \mathrm{i} t_j X_j} \\ &= \mathbb{E}\, e^{\sum_j \mathrm{i}(2\pi \ell_j \nu_n + t_j) X_j} = \widehat{f}(2\pi \ell_1 \nu_n + t_1, \dots, 2\pi \ell_k \nu_n + t_k). \end{aligned} \tag{C.6}$$

If $(\ell_1, \dots, \ell_k) = (0, \dots, 0)$, then (C.5) is trivial. If $(\ell_1, \dots, \ell_k) \neq (0, \dots, 0)$, then $|(2\pi \ell_1 \nu_n + t_1, \dots, 2\pi \ell_k \nu_n + t_k)| \to \infty$, and thus (C.6) and the Riemann–Lebesgue lemma show that the left-hand side of (C.5) tends to 0, which verifies (C.5) in this case too, since

$$\mathbb{E}\, e^{\sum_j 2\pi \mathrm{i} \ell_j U_j + \sum_j \mathrm{i} t_j X_j} = \prod_j \mathbb{E}\, e^{2\pi \mathrm{i} \ell_j U_j} \cdot \mathbb{E}\, e^{\sum_j \mathrm{i} t_j X_j} = 0$$

because $\mathbb{E}\, e^{2\pi \mathrm{i} \ell_j U_j} = 0$ when $\ell_j \neq 0$. □

Appendix D. Moments of order statistics for random p_i

Let $p_1, \dots, p_m$ be random and uniformly distributed on $\mathfrak{S}_m$ as in Section 6, and let $p_{(1)} \le \dots \le p_{(m)}$ be the order statistics. We give here, for completeness, a calculation of moments of $p_{(k)}$ by a well-known method. Note that we here follow standard convention for order statistics in probability theory and order p_i in increasing order, in contrast to Section 6 where we order them in decreasing order; thus $p_{[k]} = p_{(m+1-k)}$, which should be remembered when using the results below in Section 6.

Let $T_1, \dots, T_m$ be m independent identically distributed random variables with the exponential distribution Exp(1), and let $Z := \sum_{i=1}^m T_i$. Then $(T_1/Z, \dots, T_m/Z)$ is uniformly distributed on the simplex $\mathfrak{S}_m$, and moreover independent of Z. This means that we can take $p_i = T_i/Z$. For the corresponding order statistics we thus have $p_{(k)} = T_{(k)}/Z$; moreover $p_{(k)}$ and Z are independent and thus, for any k and ℓ,

$$\mathbb{E}(T_{(k)})^\ell = \mathbb{E}(Zp_{(k)})^\ell = \mathbb{E}(Z^\ell)\,\mathbb{E}(p_{(k)}^\ell), \tag{D.1}$$

so the moments of $p_{(k)}$ are given by

$$\mathbb{E}(p_{(k)})^\ell = \frac{\mathbb{E}(T_{(k)})^\ell}{\mathbb{E}\,Z^\ell}. \tag{D.2}$$

Furthermore, $Z := \sum_{i=1}^m T_i$ has the Gamma distribution $\Gamma(m)$ with moments

$$\mathbb{E}\,Z^\ell = \frac{\Gamma(m+\ell)}{\Gamma(m)} = m(m+1)\cdots(m+\ell-1). \tag{D.3}$$

To find the moments of the order statistics $T_{(k)}$, we regard $T_1,\dots,T_m$ as the times of the first events in m independent Poisson processes. (Alternatively, we may regard them as life-lengths of m identical radioactive atoms.) It is then well-known, and easy to see by the lack of memory in exponential distributions, that the increments (or waiting times) $V_j := T_{(j)} - T_{(j-1)}$ (with $T_{(0)} := 0$) are independent exponential random variables with $V_j \sim \mathrm{Exp}(1/(m+1-j))$. Hence $T_{(k)} = \sum_{j=1}^k V_j$ with these V_j, and moments are easily computed. In particular,

$$\mathbb{E}\,T_{(k)} = \sum_{j=1}^k \mathbb{E}\,V_j = \sum_{j=1}^k \frac{1}{m+1-j} = \sum_{i=m-k+1}^m \frac{1}{i}. \tag{D.4}$$

(This can be written as $H_m - H_{m-k}$, where $H_m = \sum_{i=1}^m 1/i$ is the m:th harmonic number.) Similarly,

$$\begin{aligned}\mathbb{E}\,T_{(k)}^2 &= \operatorname{Var} T_{(k)} + \bigl(\mathbb{E}\,T_{(k)}\bigr)^2 = \sum_{j=1}^k \operatorname{Var} V_j + \left(\sum_{j=1}^k \mathbb{E}\,V_j\right)^2 \\ &= \sum_{j=1}^k \frac{1}{(m+1-j)^2} + \left(\sum_{j=1}^k \frac{1}{m+1-j}\right)^2 \\ &= \sum_{i=m-k+1}^m \frac{1}{i^2} + \left(\sum_{i=m-k+1}^m \frac{1}{i}\right)^2. \end{aligned} \tag{D.5}$$

Hence, by (D.2)–(D.3),

$$\mathbb{E}(p_{(k)}) = \frac{\mathbb{E}\,T_{(k)}}{m} = \frac{1}{m}\sum_{i=m-k+1}^m \frac{1}{i}, \tag{D.6}$$

$$\mathbb{E}(p_{(k)})^2 = \frac{\mathbb{E}\,T_{(k)}^2}{m(m+1)} = \frac{1}{m(m+1)}\sum_{i=m-k+1}^m \frac{1}{i^2} + \frac{1}{m(m+1)}\left(\sum_{i=m-k+1}^m \frac{1}{i}\right)^2, \tag{D.7}$$

$$\operatorname{Var}(p_{(k)}) = \mathbb{E}(p_{(k)})^2 - (\mathbb{E}\,p_{(k)})^2$$

$$= \frac{1}{m(m+1)} \sum_{i=m-k+1}^{m} \frac{1}{i^2} - \frac{1}{m^2(m+1)} \left(\sum_{i=m-k+1}^{m} \frac{1}{i} \right)^2. \tag{D.8}$$

Covariances are computed similarly. If $1 \leq k \leq \ell \leq m$, then

$$\operatorname{Cov}(T_{(k)}, T_{(\ell)}) = \sum_{j=1}^{k} \operatorname{Var} V_j = \sum_{j=1}^{k} \frac{1}{(m+1-j)^2} = \sum_{i=m-k+1}^{m} \frac{1}{i^2} \tag{D.9}$$

and, calculating as in (D.5) and (D.6),

$$\operatorname{Cov}(p_{(k)}, p_{(\ell)}) = \frac{1}{m(m+1)} \sum_{i=m-k+1}^{m} \frac{1}{i^2} - \frac{1}{m^2(m+1)} \sum_{i=m-k+1}^{m} \frac{1}{i} \sum_{i=m-\ell+1}^{m} \frac{1}{i}. \tag{D.10}$$

Example D.1. For $m = 3$, the covariance matrix of $(p_{[1]}, p_{[2]}, p_{[3]}) = (p_{(3)}, p_{(2)}, p_{(1)})$ is

$$\frac{1}{648} \begin{pmatrix} 13 & -8 & -5 \\ -8 & 7 & 1 \\ -5 & 1 & 4 \end{pmatrix}.$$

References

[1] M. L. Balinski & H. P. Young, *Fair Representation.* 2nd ed., Brookings Institution Press, Washington, D.C., 2001.

[2] P. Billingsley, *Convergence of Probability Measures.* Wiley, New York, 1968.

[3] N. H. Bingham, C. M. Goldie & J. L. Teugels, *Regular Variation.* Cambridge Univ. Press, Cambridge, 1987.

[4] A. M. Carstairs, *A Short History of Electoral Systems in Western Europe.* George Allen & Unwin, London, 1980.

[5] V. D'Hondt, *Système pratique et raisonné de représentation proportionnelle.* Bruxelles, 1882.

[6] M. Drton & U. Schwingenschlögl, Asymptotic seat bias formulas. *Metrika* **62** (2005), no. 1, 23–31.

[7] David M. Farrell, *Electoral Systems.* 2nd ed., Palgrave Macmillan, Basingstoke, 2011.

[8] N. Gaffke & F. Pukelsheim, Divisor methods for proportional representation systems: An optimization approach to vector and matrix apportionment problems. *Mathematical Social Sciences* **56** (2008), 166–184.

[9] L. Grafakos, *Classical and Modern Fourier Analysis.* Pearson, Upper Saddle River, NJ, 2004.

[10] G. R. Grimmett, European Apportionment via the Cambridge Compromise. *Mathematical Social Sciences* **63** (2012), 68–73.

[11] A. Gut, *Probability: A Graduate Course.* Springer, New York, 2005.

[12] L. Heinrich, F. Pukelsheim & U. Schwingenschlögl, Sainte-Laguë's chi-square divergence for the rounding of probabilities and its convergence to a stable law. *Statistics & Decisions* **22** (2004), 43–59.

[13] L. Heinrich, F. Pukelsheim & U. Schwingenschlögl, On stationary multiplier methods for the rounding of probabilities and the limiting law of the Sainte-Laguë divergence. *Statistics & Decisions* **23** (2005), 117–129.

[14] L. Heinrich & U. Schwingenschlögl, Goodness-of-fit criteria for the Adams and Jefferson rounding methods and their limiting laws. *Metrika* **64** (2006), no. 2, 191–207.

[15] E. V. Huntington, The apportionment of representatives in Congress. *Trans. Amer. Math. Soc.* **30** (1928), no. 1, 85–110.

[16] S. Janson, Another note on the Droop quota and rounding. *Voting matters*, **29** (2011), 32–34.

[17] S. Janson & S. Linusson, The probability of the Alabama paradox. *J. Appl. Probab.*, **49** (2012), to appear.

[18] Klaus Kopfermann, *Mathematische Aspekte der Wahlverfahren.* Wissenschaftsverlag, Mannheim, 1991.

[19] P. Leutgäb & F. Pukelsheim, List apparentements in local elections – a lottery. *Homo Oeconomicus* **26** (2009), no. 3/4, 489–500.

[20] H. F. Niemeyer & A. C. Niemeyer, Apportionment methods. *Math. Social Sci.* **56** (2008), no. 2, 240–253.

[21] H. Niemeyer & G. Wolf, Über einige mathematische Aspekte bei Wahlverfahren. *Z. Angew. Math. Mech.* **64** (1984), no. 5, 340–343.

[22] K.-F. Oelbermann, F. Pukelsheim & A. Palomares, The 2009 European Parliament elections: From votes to seats in 27 ways. *European Electoral Studies* **5**, (2010), no. 1, 148–182.

[23] G. Pólya, Über die Verteilungssysteme der Proportionalwahl. *Zeitschrift für schweizerische Statistik und Volkswirtschaft* **54** (1918), 363–387.

[24] G. Pólya, Sur la représentation proportionnelle en matière électorale. *L'Enseignement Mathématique* **20** (1918), 355–379.

[25] G. Pólya, Proportionalwahl und Wahrscheinlichkeitsrechnung. *Zeitschrift für die ges. Staatswissenschaft* **74** (1919), 297–322.

[26] A. Sainte-Laguë, La représentation proportionnelle et la méthode des moindres carrés. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, **151** (1910), 377–378.

[27] A. Sainte-Laguë, La représentation proportionnelle et la méthode des moindres carrés. *Ann. Sci. École Norm. Sup.* (3) **27** (1910), 529–542.

[28] K. Schuster, F. Pukelsheim, M. Drton & N. R. Draper, Seat biases of apportionment methods for proportional representation. *Electoral Studies* **22** (2003), 651–676.

[29] U. Schwingenschlögl, Seat biases of apportionment methods under general distributional assumptions. *Appl. Math. Lett.* **21** (2008), no. 1, 1–3.

[30] U. Schwingenschlögl & M. Drton, Seat allocation distributions and seat biases of stationary apportionment methods for proportional representation. *Metrika* **60** (2004), no. 2, 191–202.

[31] U. Schwingenschlögl & M. Drton, Seat excess variances of apportionment methods for proportional representation. *Statist. Probab. Lett.* **76** (2006), no. 16, 1723–1730.

[32] H. Weyl, Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.* **77** (1916), no. 3, 313–352.

Examples of official regulations (accessed 21 October 2011)

[33] Belgium: *Loi électorale communale du 4 août 1932.* `http://elections2006.wallonie.be/apps/spip/IMG/pdf/loi_electorale_1932-2.pdf`

[34] Estonia: *Riigikogu Election Act.* `http://www.vvk.ee/public/dok/RKseadus_eng_2010.pdf`

[35] Germany: *Bundeswahlgesetz.* `www.bundestag.de/dokumente/rechtsgrundlagen/bwahlg_pdf.pdf`

[36] Macau: *Lei Eleitoral n.º 3/2001.* `http://bo.io.gov.mo/bo/i/2001/10/lei03.asp`

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: svante.janson@math.uu.se

URL: http://www.math.uu.se/~svante/