The fate of the square root law for correlated voting

Voting Power in Practice

London School of Economics
Symposium sponsored by the Leverhulme Trust

Werner Kirsch and Jessica Langner
Fakultät für Mathematik and Informatik
FernUniversität Hagen
talk based on:


talk based on:


strongly influenced by:

Some Two-Tier Voting Systems (Councils)

The fate of the square root law for correlated voting

London School of Economics, March 2011
In a Two-Tier Voting Systems a union is formed by member states (or member organizations). The voters in the member states are represented by a delegate. The delegates of the states are the member of a Council.
In a **Two-Tier Voting System** a union is formed by member states (or member organizations).

The voters in the member states are represented by a delegate. The delegates of the states are the member of a **Council**.

The delegates in the council are given a certain voting weight depending on the size of the country they represent (or its economic strength, or ...).

The delegate’s vote cannot be split.
What is a ‘fair’ distribution of votes in a council?

We regard the process of decision making in a council as a two-step voting system (two-tier system). In a first step all voters in the different states decide upon the voting of their respective representative in the council (one person, one vote). The delegates cast their votes in the council according to the majority of voters in her/his country. The voting weight of a delegate depends on the population of his/her country. A reasonable criterion for ‘fair’ voting weights:

The decision of the council should agree with the public vote!
We regard the process of decision making in a council as a two-step voting system (two-tier system).

In a first step all voters in the different states decide upon the voting of their respective representative in the council (one person, one vote).

The delegates cast their votes in the council according to the majority of voters in her/his country. The voting weight of a delegate depends on the population of his/her country.
We regard the process of decision making in a council as a two-step voting system (two-tier system).

In a first step all voters in the different states decide upon the voting of their respective representative in the council (one person, one vote).

The delegates cast their votes in the council according to the majority of voters in her/his country. The voting weight of a delegate depends on the population of his/her country.

A reasonable criterion for ‘fair’ voting weights:
The decision of the council should agree with the public vote!
In the year 2000 presidential election in the USA Al Gore won the majority in the public vote but Georg W. Bush was elected by the Electoral College.
In the year 2000 presidential election in the USA Al Gore won the majority in the public vote but Georg W. Bush was elected by the Electoral College.

No distribution of voting weights can guarantee that the public vote and the council vote coincide.
In the year 2000 presidential election in the USA Al Gore won the majority in the public vote but Georg W. Bush was elected by the Electoral College.

No distribution of voting weights can guarantee that the public vote and the council vote coincide.

The best we can do is to choose the voting weights in such a way that *in most cases* public vote and council vote agree.
A Mathematical Formulation

States and Voters

- There are $M$ states, labeled by Greek characters $\nu, \kappa, \ldots$.
- There are $N_\nu$ voters in the state $\nu$. The votes of the citizens in state $\nu$ are denoted by $X_{\nu i}$ with $i = 1, \ldots, N_\nu$ and

$$X_{\nu i} = \begin{cases} 
1, & \text{for ‘yes’;} \\
-1, & \text{for ‘no’}.
\end{cases}$$
A Mathematical Formulation

States and Voters

- There are $M$ states, labeled by Greek characters $\nu, \kappa, \ldots$.
- There are $N_\nu$ voters in the state $\nu$. The votes of the citizens in state $\nu$ are denoted by $X_{\nu i}$ with $i = 1, \ldots, N_\nu$ and

$$X_{\nu i} = \begin{cases} 
1, & \text{for ‘yes’;} \\
-1, & \text{for ‘no’}.
\end{cases}$$

- The voting result in state $\nu$ is represented by

$$S_\nu = \sum_{i=1}^{N_\nu} X_{\nu i}$$

If $S_\nu > 0$ the majority in state $\nu$ votes affirmatively.

- The popular vote in the union is given by:

$$P = \sum_{\nu=1}^{M} S_\nu = \sum_{\nu=1}^{M} \sum_{i=1}^{N_\nu} X_{\nu i}$$

The fate of the square root law for correlated voting

London School of Economics, March 2011
We set
\[ \chi(x) := \begin{cases} 
  1, & \text{for } x > 0; \\
  -1, & \text{otherwise.}
\end{cases} \]

Thus
\[ \chi_\nu := \chi(S_\nu) = \begin{cases} 
  1, & \text{if the majority in state } \nu \text{ votes ‘yes’;} \\
  -1, & \text{if not.}
\end{cases} \]
We set
\[ \chi(x) := \begin{cases} 1, & \text{for } x > 0; \\ -1, & \text{otherwise}. \end{cases} \]

Thus
\[ \chi_\nu := \chi(S_\nu) = \begin{cases} 1, & \text{if the majority in state } \nu \text{ votes ‘yes’}; \\ -1, & \text{if not}. \end{cases} \]

If the state \( \nu \) has voting weight \( w_\nu \), then the vote in the council will be represented by
\[ C = \sum_{\nu=1}^{M} w_\nu \chi(S_\nu) \]
We set

\[ \chi(x) := \begin{cases} 1, & \text{for } x > 0; \\ -1, & \text{otherwise.} \end{cases} \]

Thus

\[ \chi_\nu := \chi(S_\nu) = \begin{cases} 1, & \text{if the majority in state } \nu \text{ votes ‘yes’;} \\ -1, & \text{if not.} \end{cases} \]

If the state \( \nu \) has voting weight \( w_\nu \), then the vote in the council will be represented by

\[ C = \sum_{\nu=1}^{M} w_\nu \chi(S_\nu) \]

We define the democracy deficit by

\[ \Delta := |C - P| \]

The democracy deficit \( \Delta(w_1, \ldots, w_M) \) is a function of the weights \( w_\nu, \nu = 1 \ldots, M \).
Choose the weights $w_\nu$ in such a way that $\Delta(w_1, \ldots, w_M)$ is minimal!

More precisely we want the average square error $E(\Delta^2) = E\left(\sum_{\nu=1}^{M} w_\nu \chi(N_\nu \sum_{i=1}^{N_\nu} X_\nu i) - \sum_{\nu=1}^{M} N_\nu \sum_{i=1}^{N_\nu} X_\nu i\right)^2$ to be as small as possible.

But . . . what is $E(=\text{average})$ supposed to mean?
Choose the weights $w_\nu$ in such a way that $\Delta(w_1, \ldots, w_M)$ is minimal!

More precisely, we want the average square error

$$\mathbb{E}(\Delta^2) = \mathbb{E} \left( \left( \sum_{\nu=1}^{M} w_\nu \chi \left( \sum_{i=1}^{N_\nu} x_{\nu i} \right) - \sum_{\nu=1}^{M} \sum_{i=1}^{N_\nu} x_{\nu i} \right)^2 \right)$$

to be as small as possible.
Choose the weights $w_\nu$ in such a way that $\Delta(w_1, \ldots, w_M)$ is minimal!

More precisely we want the average square error

$$\mathbb{E}(\Delta^2) = \mathbb{E}\left( \left( \sum_{\nu=1}^{M} w_\nu \chi \left( \sum_{i=1}^{N_\nu} X_{\nu i} \right) - \sum_{\nu=1}^{M} \sum_{i=1}^{N_\nu} X_{\nu i} \right)^2 \right)$$

to be as small as possible.

? 

But . . . what is $E(=\text{average})$ supposed to mean?

The fate of the square root law for correlated voting

London School of Economics, March 2011
Choose the weights \( w_\nu \) in such a way that \( \Delta(w_1, \ldots, w_M) \) is minimal!
More precisely we want the average square error
\[
\mathbb{E}\left( \Delta^2 \right) = \mathbb{E}\left( \left( \sum_{\nu=1}^{M} w_\nu \chi\left( \sum_{i=1}^{N_\nu} X_{\nu i} \right) - \sum_{\nu=1}^{M} \sum_{i=1}^{N_\nu} X_{\nu i} \right)^2 \right)
\]
to be as small as possible.

But . . .

... what is \( \mathbb{E} \) (\( = \)average) supposed to mean?
The voting system, represented by $X_1, \ldots, X_N$, is fed with proposals in a completely random way. The voters react to the (random) proposal in a deterministic rational way.

The randomness of the proposals induce a probability measure $P$ on the space $\{-1, +1\}^N$, the possible voting outcomes. A proposal $\omega$ and its counter-proposal $\neg \omega$ should have the same probability. The rationality of the voting implies that $X_v(\neg \omega) = -X_v(\omega)$. Consequently, the probability $P$ is invariant under inversion $(X_1, \ldots, X_N) \rightarrow (-X_1, \ldots, -X_N)$.

In particular:

$$P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$$

The fate of the square root law for correlated voting
The voting system, represented by $X_1, \ldots, X_N$, is fed with proposals in a completely random way. The voters react to the (random) proposal in a deterministic rational way.

The randomness of the proposals induce a probability measure $\mathbb{P}$ on the space $\{-1, +1\}^N$, the possible voting outcomes. A proposal $\omega$ and its counter-proposal $\neg \omega$ should have the same probability. The rationality of the voting implies that $X_v(\neg \omega) = -X_v(\omega)$. Consequently, the probability $\mathbb{P}$ is invariant under inversion

$$(X_1, \ldots, X_N) \longrightarrow (-X_1, \ldots, -X_N)$$

In particular:

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$$
A measure $\mathbb{P}$ on $\{-1, 1\}^N$ is called a **voting measure** if

$$\mathbb{P}(X_1 = \xi_1, \ldots, X_N = \xi_N) = \mathbb{P}(X_1 = -\xi_1, \ldots, X_N = -\xi_N)$$

**Example: Independent Voters**

If the voting behavior $X_i$ of the voter $i$ is independent from the other voters, then $\mathbb{P}$ is just a product measure $\mathbb{P}_0$ with $\mathbb{P}_0 = \frac{1}{2}(\delta_{-1} + 1)$, that is

$$\mathbb{P}(X_1 = \xi_1, \ldots, X_N = \xi_N) = \frac{1}{2^N}.$$
A measure $\mathbb{P}$ on $\{-1, 1\}^N$ is called a voting measure if

$$\mathbb{P}(X_1 = \xi_1, \ldots, X_N = \xi_N) = \mathbb{P}(X_1 = -\xi_1, \ldots, X_N = -\xi_N)$$

Example: Independent Voters

If the voting behavior $X_i$ of the voter $i$ is independent from the other voters, then $\mathbb{P}$ is just a product measure $P_0$ with

$$P_0 = \frac{1}{2}(\delta_{-1} + \delta_{+1})$$

that is

$$\mathbb{P}(X_1 = \xi_1, \ldots, X_N = \xi_N) = \frac{1}{2^N}$$
We assume now that $X_{\nu i}$ and $X_{\kappa j}$ are independent for $\nu \neq \kappa$.

We keep the notation $S_\nu = \sum_i X_{\nu i}$ and $\chi_\nu = \pm 1$ for $S_\nu > 0$ (resp. $S_\nu \leq 0$).

We want to minimize the function

$$\mathcal{D}(w_1, \ldots, w_M) = \mathbb{E}(\Delta(w_1, \ldots, w_M)^2)$$

$$= \sum_{\nu, \kappa = 1}^M \left( w_\nu w_\kappa \mathbb{E}(\chi_\nu \chi_\kappa) - 2w_\nu \mathbb{E}(\chi_\nu S_\kappa) + \mathbb{E}(S_\nu S_\kappa) \right)$$

The fate of the square root law for correlated voting

London School of Economics, March 2011
We assume now that $X_{\nu i}$ and $X_{\kappa j}$ are independent for $\nu \neq \kappa$

We keep the notation $S_{\nu} = \sum_i X_{\nu i}$

and $\chi_{\nu} = \pm 1$ for $S_{\nu} > 0$ (resp. $S_{\nu} \leq 0$).

We want to minimize the function

$$D(w_1, \ldots, w_M) = \mathbb{E}(\Delta(w_1, \ldots, w_M)^2)$$

$$= \sum_{\nu, \kappa=1}^{M} \left( w_{\nu} w_{\kappa} \mathbb{E}(\chi_{\nu} \chi_{\kappa}) - 2 w_{\nu} \mathbb{E}(\chi_{\nu} S_{\kappa}) + \mathbb{E}(S_{\nu} S_{\kappa}) \right)$$

We have (in general) $\chi_{\nu}^2 = 1$ and $\chi_{\nu} S_{\nu} = |S_{\nu}|$.

Moreover, due to independence, we know for $\nu \neq \kappa$

$$\mathbb{E}(\chi_{\nu} \chi_{\kappa}) = 0$$

$$\mathbb{E}(\chi_{\nu} S_{\kappa}) = 0$$

$$\mathbb{E}(S_{\nu} S_{\kappa}) = 0$$
Independent States

The $\mathcal{D}$ simplifies to:

$$\mathcal{D}(w_1, \ldots, w_M) = \sum_{\nu=1}^{M} \left( w_{\nu}^2 - 2w_{\nu} \mathbb{E}(|S_{\nu}|) + \mathbb{E}(S_{\nu}^2) \right).$$

This is minimized by:

$$g_{\nu} := \mathbb{E}(|S_{\nu}|)$$

The quantity $\mathbb{E}(|S_{\nu}|)$ describes the margin of the voting outcome.
Independent States

The $\mathcal{D}$ simplifies to:

$$\mathcal{D}(w_1, \ldots, w_M) = \sum_{\nu=1}^{M} \left( w_{\nu}^2 - 2w_{\nu} \mathbb{E}(|S_{\nu}|) + \mathbb{E}(S_{\nu}^2) \right).$$

This is minimized by:

$$g_{\nu} := \mathbb{E}(|S_{\nu}|)$$

The quantity $\mathbb{E}(|S_{\nu}|)$ describes the margin of the voting outcome.

Democracy Deficit

The minimal value of $\mathcal{D}$ is given by:

$$\mathcal{D}_0 := \mathcal{D}(g_1, \ldots, g_M) = \sum_{\nu=1}^{M} \left( \mathbb{E}(|S_{\nu}|^2) - \mathbb{E}(|S_{\nu}|)^2 \right).$$
Now we assume that all voters are independent.

Independent Voters

For independent voters we have

$$g_{\nu} = E(|S_{\nu}|) \approx \sqrt{\frac{2}{\pi}} \sqrt{N}$$

This is the square root law by Penrose.

Democracy Deficit

The minimal (averaged) democracy deficit is given by:

$$D_0 = D(g_1, \ldots, g_M) \approx \pi - 2\pi N$$

The averaged democracy deficit per voter converges to 0:

$$E((\Delta_0 N)^2) \leq C N$$

The fate of the square root law for correlated voting

London School of Economics, March 2011
Now we assume that all voters are independent.

**Independent Voters**

For independent voters we have

$$g_\nu = \mathbb{E}(|S_\nu|) \approx \sqrt{\frac{2}{\pi}} \sqrt{N}$$

This is the square root law by Penrose.
Now we assume that all voters are independent.

**Independent Voters**

For independent voters we have

\[ g_\nu = \mathbb{E}(|S_\nu|) \approx \sqrt{\frac{2}{\pi}} \sqrt{N} \]

This is the *square root law* by Penrose.

**Democracy Deficit**

The minimal (averaged) democracy deficit is given by:

\[ D_0 = D(g_1, \ldots, g_M) \approx \frac{\pi - 2}{\pi} N \]

The averaged democracy deficit per voter converges to 0:

\[ \mathbb{E}\left( \left( \frac{\Delta_0}{N} \right)^2 \right) \leq \frac{C}{N} \]

The fate of the square root law for correlated voting
In this model the voters are influenced by a mainstream opinion, e.g. a common believe due to the country’s tradition or the influence of opinion makers.

For a given proposal $\omega$ we model this ‘common believe’ by a value $\zeta \in [-1, 1]$ which depends on the proposal at hand. $\zeta = 1$ means there is such a strong common believe in favor of the proposal that all voters will vote ‘yes’, $\zeta = -1$ means all voters will vote ‘no’. In general, $\zeta$ denotes the expected outcome of the voting, i.e. $E(X_i)$. The individual voting results $X_i$ fluctuate around this value randomly.
The Collective Bias Model

In this model the voters are influenced by a mainstream opinion, e.g. a common believe due to the country’s tradition or the influence of opinion makers.

For a given proposal $\omega$ we model this ‘common believe’ by a value $\zeta \in [-1, 1]$ which depends on the proposal at hand.

The value $\zeta = 1$ means there is such a strong common believe in favor of the proposal that all voters will vote ‘yes’, $\zeta = -1$ means all voters will vote ‘no’.
In this model the voters are influenced by a mainstream opinion, e. g. a common believe due to the country’s tradition or the influence of opinion makers.

For a given proposal $\omega$ we model this ‘common believe’ by a value $\zeta \in [-1, 1]$ which depends on the proposal at hand.

The value $\zeta = 1$ means there is such a strong common believe in favor of the proposal that all voters will vote ‘yes’, $\zeta = -1$ means all voters will vote ‘no’.

In general, $\zeta$ denotes the expected outcome of the voting, i. e. $E(X_i)$. The individual voting results $X_i$ fluctuate around this value randomly.
For $\zeta \in [-1, 1]$ let $P_\zeta$ be the probability measure on $\{-1, 1\}$ with $P_\zeta(1) = \frac{1}{2}(1 + \zeta)$ and $P_\zeta(-1) = \frac{1}{2}(1 - \zeta)$. Then $E_\zeta(X_i) = \zeta$.

We define $P_\zeta$ on $\{-1, 1\}^N$ by

$$P_\zeta(X_1 = a_1, X_2 = a_2, \ldots, X_N = a_N) = P_\zeta(a_1) \cdot P_\zeta(a_2) \cdot \ldots \cdot P_\zeta(a_N)$$
CBM (mathematical definition)

For $\zeta \in [-1, 1]$ let $P_\zeta$ be the probability measure on $\{-1, 1\}$ with $P_\zeta(1) = \frac{1}{2}(1 + \zeta)$ and $P_\zeta(-1) = \frac{1}{2}(1 - \zeta)$ . Then $E_\zeta(X_i) = \zeta$.

We define $\mathcal{P}_\zeta$ on $\{-1, 1\}^N$ by

$$\mathcal{P}_\zeta(X_1 = a_1, X_2 = a_2, \ldots, X_N = a_N) = P_\zeta(a_1) \cdot P_\zeta(a_2) \cdot \ldots \cdot P_\zeta(a_N)$$

Let $\mu$ be a measure on $[-1, 1]$ with $\mu([a, b])$ giving the probability that $\zeta$ has a value between $a$ and $b$.

We define the CBM-measure by

$$\mathbb{P}_\mu(X_1 = a_1, X_2 = a_2, \ldots, X_N = a_N) = \int_{-1}^{1} \mathcal{P}_\zeta(X_1 = a_1, X_2 = a_2, \ldots, X_N = a_N) \, d\mu(\zeta)$$
Note, that $P_\zeta$ is not a voting measure (unless $\zeta = 0$).
Note, that $\mathcal{P}_\zeta$ is not a voting measure (unless $\zeta = 0$).

However $\mathcal{P}_\mu$ is a voting measure if $\mu$ is invariant under sign change. We call $\mu$ the bias measure.
Note, that $\mathcal{P}_\zeta$ is *not* a voting measure (unless $\zeta = 0$).

However $\mathbb{P}_\mu$ is a voting measure if $\mu$ is invariant under sign change. We call $\mu$ the *bias measure*.

If the measure $\mu$ is concentrated in 0, then $\mathbb{P}_\mu$ makes the voting results $X_i$ independent, thus we are in the case of independent voters.
Note, that $P_\zeta$ is not a voting measure (unless $\zeta = 0$).

However $P_\mu$ is a voting measure if $\mu$ is invariant under sign change. We call $\mu$ the bias measure.

If the measure $\mu$ is concentrated in 0, then $P_\mu$ makes the voting results $X_i$ independent, thus we are in the case of independent voters.

If $\mu$ is the uniform distribution on $[-1,1]$, then the corresponding measure was already considered by Straffin (1977) when he established an intimate connection to of this model to the Shapley-Shubik index.
Note, that $P_\zeta$ is not a voting measure (unless $\zeta = 0$).

However $P_\mu$ is a voting measure if $\mu$ is invariant under sign change. We call $\mu$ the bias measure.

If the measure $\mu$ is concentrated in 0, then $P_\mu$ makes the voting results $X_i$ independent, thus we are in the case of independent voters.

If $\mu$ is the uniform distribution on $[-1, 1]$, then the corresponding measure was already considered by Straffin (1977) when he established an intimate connection to of this model to the Shapley-Shubik index.

In a similar way, the Penrose-Banzhaf measure is connected with the model of independent voters.
Optimal weights

For the CBM we have

\[ g_\nu = \mathbb{E}(|S_\nu|) \approx \mu_1 N \]

where \( \mu_1 = \int |\zeta| \, d\mu(\zeta) \).
Optimal weights

For the CBM we have
\[ g_\nu = \mathbb{E}(|S_\nu|) \approx \mu_1 N \]
where \( \mu_1 = \int |\zeta| d\mu(\zeta) \).

Democracy Deficit for the CBM

The minimal (averaged) democracy deficit is given by:
\[ D_0 \approx (\mu_2 - \mu_1^2) N^2 \]
with \( \mu_2 = \int |\zeta|^2 d\mu(\zeta) \).
The averaged democracy deficit per voter converges to a nonzero limit:
\[ \mathbb{E}\left(\left(\frac{\Delta_0}{N}\right)^2\right) \rightarrow \mu_2 - \mu_1^2 \]
The **Curie-Weiss model** is a model of cooperative behavior from statistical physics. It models ferromagnetic system in which elementary magnets prefer to align.

\[ H(X_1, \ldots, X_N) = -\frac{1}{N} \sum_{i=1}^{N} X_i \]

This is the energy function for the spin configuration (=voting outcome) \( X_1, \ldots, X_N \). We use this to define probability measures

\[ P_\beta(X_1, \ldots, X_N) = \frac{e^{-\beta H(X_1, \ldots, X_N)}}{Z} \]

The parameter \( \beta \) measures the strength of the interaction between the voters.
The Curie-Weiss model is a model of cooperative behavior from statistical physics. It models ferromagnetic system in which elementary magnets prefer to align.

Curie Weiss Measure

\[ H(X_1, \ldots, X_N) = -\frac{1}{N} \left( \sum_{i=1}^{N} X_i \right)^2 \]

This is the energy function for the spin configuration (=voting outcome) \( X_1, \ldots, X_N \). We use this to define probability measures

\[ P_\beta(X_1, \ldots, X_N) = \frac{e^{-\beta H(X_1, \ldots, X_N)}}{Z} \]

The parameter \( \beta \) measures the strength of the interaction between the voters.

The fate of the square root law for correlated voting

London School of Economics, March 2011
Optimal Weights for the Curie-Weiss Model

\[
g_{\nu} = \mathbb{E}(|S_{\nu}|) \approx \begin{cases} 
C_1(\beta) \sqrt{N_{\nu}}, & \text{for } \beta < 1; \\
C_2 N_{\nu}^{\frac{3}{4}}, & \text{for } \beta = 1; \\
C_3(\beta) N_{\nu}, & \text{for } \beta > 1.
\end{cases}
\]
So far we have always assumed that voter in different states act independently.

In this section we consider the case of collective behavior across country borders.
So far we have always assumed that voter in different states act independently.

In this section we consider the case of collective behavior across country borders.

We assume that all voters act according to the Collective Bias measure $P_{\mu}$.

This means there is a common believe, expressed through the measure $\mu$, for all voters in the union.
We have again to minimize the function

\[ D(w_1, \ldots, w_M) = \mathbb{E}(\Delta(w_1, \ldots, w_M)^2) \]

\[ = \sum_{\nu,\kappa=1}^{M} (w_\nu w_\kappa \mathbb{E}(\chi_\nu \chi_\kappa) - 2w_\nu \mathbb{E}(\chi_\nu S_\kappa) + \mathbb{E}(S_\nu S_\kappa)) \]

But this time the function does not simplify immediately!
We have again to minimize the function

\[
D(w_1, \ldots, w_M) = \mathbb{E}(\Delta(w_1, \ldots, w_M)^2) = \sum_{\nu, \kappa = 1}^{M} \left( w_\nu w_\kappa \mathbb{E}(\chi_\nu \chi_\kappa) - 2w_\nu \mathbb{E}(\chi_\nu S_\kappa) + \mathbb{E}(S_\nu S_\kappa) \right)
\]

But this time the function does not simplify immediately!

However, for large \( N_\nu \) and \( N_\kappa \) we have for any \( \nu, \kappa \) :

\[
\mathbb{E}_\mu(\chi_\nu \chi_\kappa) \approx 1 \\
\mathbb{E}_\mu(\chi_\nu S_\kappa) \approx \mathbb{E}_\mu(|S_\kappa|) \approx \mu_1 N_\kappa
\]

and

\[
\mathbb{E}(S_\nu S_\kappa) \approx \mu_2 N_\nu N_\kappa.
\]
For large $N$ we obtain:

$$D(w_1, \ldots, w_M) \approx \sum_{\nu,\kappa=1}^{M} w_{\nu} w_{\kappa} + 2 \sum_{\nu=1}^{M} w_{\nu} \sum_{\kappa=1}^{M} \mu_1 N_{\kappa} + \sum_{\nu,\kappa=1}^{M} \mu_2 N_{\nu} N_{\kappa}$$

$$= (\sum_{\nu=1}^{M} w_{\nu})^2 - 2\mu_1 (\sum_{\nu=1}^{M} w_{\nu}) N + \mu_2 N^2$$

This last expression depends only on the sum $W = \sum_{\nu=1}^{M} w_{\nu}$ of the voting weights and not on the individual weights $w_{\nu}$.
For large $N$ we obtain:

$$\mathcal{D}(w_1, \ldots, w_M) \approx \sum_{\nu, \kappa=1}^{M} w_\nu w_\kappa + 2 \sum_{\nu=1}^{M} w_\nu \sum_{\kappa=1}^{M} \mu_1 N_\kappa + \sum_{\nu, \kappa=1}^{M} \mu_2 N_\nu N_\kappa$$

$$= \left( \sum_{\nu=1}^{M} w_\nu \right)^2 - 2\mu_1 \left( \sum_{\nu=1}^{M} w_\nu \right) N + \mu_2 N^2$$

This last expression depends only on the sum $W = \sum_{\nu=1}^{M} w_\nu$ of the voting weights and not on the individual weights $w_\nu$. For large $N$, the asymptotic value of $\mathcal{D}$ does not depend on the way the weights are distributed among the member states of the union.
If the voters act independently or almost independently, the square root law applies.

If the voters inside a country are strongly correlated to each other, but not to voters in other countries, then proportional representation is optimal.

If all voters are strongly correlated (also across borders), any distribution of voting weights is as good as any other one.