DAV / DGVFM
Jahrestagung
2025

DAV
Deutsche Aktuarvereinigung e.V.

DGVFM
Deutsche Gesellschaft für Versicherungs- und Finanzmathematik e.V.

Fatima ezzahra Kherraz, Dr. Ulrich Riegel

Burning Cost or Pareto?

Jahrestagung, 29. April 2025

Agenda

Excess of loss per risk (per risk XL)

We consider the annual loss $S = \sum_{i=1}^{N} X_i$ of an insurance portfolio. X_i denotes the size of the i-th individual loss.

Excess of loss per risk (per risk XL)

We consider the annual loss $S = \sum_{i=1}^{N} X_i$ of an insurance portfolio. X_i denotes the size of the i-th individual loss.

- A *per risk XL* has a *deductible* D and a *cover* C

Excess of loss per risk (per risk XL)

We consider the annual loss $S = \sum_{i=1}^{N} X_i$ of an insurance portfolio. X_i denotes the size of the i-th individual loss.

- A *per risk XL* has a *deductible* D and a *cover* C
- Basic variant:

Reinsurer: $$\hat{S} = \sum_{i=1}^{N} \min(C, (X_i - D)^+) =: \sum_{i=1}^{N} \hat{X}_i, \qquad (x^+ := \max(x, 0))$$

Cedent: $$\tilde{S} = S - \hat{S} = \sum_{i=1}^{N} (X_i - \hat{X}_i).$$

Excess of loss per risk (per risk XL)

We consider the annual loss $S = \sum_{i=1}^{N} X_i$ of an insurance portfolio. X_i denotes the size of the i-th individual loss.

- A *per risk XL* has a *deductible D* and a *cover C*
- Basic variant:

  Reinsurer: $$\hat{S} = \sum_{i=1}^{N} \min(C, (X_i - D)^+) =: \sum_{i=1}^{N} \hat{X}_i, \qquad (x^+ := \max(x, 0))$$

  Cedent: $$\tilde{S} = S - \hat{S} = \sum_{i=1}^{N} (X_i - \hat{X}_i).$$

- Most per risk XLs have an annual aggregate limit (AAL) for the reinsurer. Sometimes, there is an annual aggregate deductible (AAD) of the cedent.

Excess of loss per risk (per risk XL)

We consider the annual loss $S = \sum_{i=1}^{N} X_i$ of an insurance portfolio. X_i denotes the size of the i-th individual loss.

- A *per risk XL* has a *deductible* D and a *cover* C
- Basic variant:

  Reinsurer: $$\hat{S} = \sum_{i=1}^{N} \min(C, (X_i - D)^+) =: \sum_{i=1}^{N} \hat{X}_i, \qquad (x^+ := \max(x, 0))$$

  Cedent: $$\tilde{S} = S - \hat{S} = \sum_{i=1}^{N} (X_i - \hat{X}_i).$$
- Most per risk XLs have an annual aggregate limit (AAL) for the reinsurer. Sometimes, there is an annual aggregate deductible (AAD) of the cedent.
- C xs D is referred to as a *layer*. A reinsurance program often consists of several layers C_i xs D_i, such that $D_{i+1} = C_i + D_i$.

Excess of loss per risk (per risk XL)

We consider the annual loss $S = \sum_{i=1}^{N} X_i$ of an insurance portfolio. X_i denotes the size of the i-th individual loss.

- A *per risk XL* has a *deductible D* and a *cover C*
- Basic variant:

  Reinsurer: $$\hat{S} = \sum_{i=1}^{N} \min(C, (X_i - D)^+) =: \sum_{i=1}^{N} \hat{X}_i, \qquad (x^+ := \max(x, 0))$$

  Cedent: $$\tilde{S} = S - \hat{S} = \sum_{i=1}^{N} (X_i - \hat{X}_i).$$

- Most per risk XLs have an annual aggregate limit (AAL) for the reinsurer. Sometimes, there is an annual aggregate deductible (AAD) of the cedent.
- C xs D is referred to as a *layer*. A reinsurance program often consists of several layers C_i xs D_i, such that $D_{i+1} = C_i + D_i$.
- Purpose of a per risk XL: Protection against large losses.

Experience rating

Basic idea:

Estimate the expected loss of a per risk XL in the prospective treaty year based on the individual loss experience of the cedent in a certain *observation period*.

Experience rating

Basic idea:

Estimate the expected loss of a per risk XL in the prospective treaty year based on the individual loss experience of the cedent in a certain *observation period*.

Starting point:

- Observation period: Accident years $1, \ldots, n$
- Current year: $n+1$
- Prospective treaty year: $q = n+2$

Experience rating

Basic idea:

Estimate the expected loss of a per risk XL in the prospective treaty year based on the individual loss experience of the cedent in a certain *observation period*.

Starting point:

- Observation period: Accident years $1, \ldots, n$
- Current year: $n+1$
- Prospective treaty year: $q = n+2$
- Volumes $v_1, \ldots, v_q$ available (e.g., premiums, sums insured, vehicle counts)

Experience rating

Basic idea:

Estimate the expected loss of a per risk XL in the prospective treaty year based on the individual loss experience of the cedent in a certain *observation period*.

Starting point:

- Observation period: Accident years $1, \ldots, n$
- Current year: $n+1$
- Prospective treaty year: $q = n+2$
- Volumes $v_1, \ldots, v_q$ available (e.g., premiums, sums insured, vehicle counts)
- $X_{i,1}, \ldots, X_{i,N_i}$ observed individual losses in the year $i \in \{1, \ldots, n\}$

Experience rating

Basic idea:

Estimate the expected loss of a per risk XL in the prospective treaty year based on the individual loss experience of the cedent in a certain *observation period*.

Starting point:

- Observation period: Accident years $1, \ldots, n$
- Current year: $n+1$
- Prospective treaty year: $q = n+2$
- Volumes $v_1, \ldots, v_q$ available (e.g., premiums, sums insured, vehicle counts)
- $X_{i,1}, \ldots, X_{i,N_i}$ observed individual losses in the year $i \in \{1, \ldots, n\}$

We assume that the volumes and the losses have been *indexed* to the cost level of the prospective treaty year q.

Burning Cost (BC)

An important step in the pricing of the layer C xs D is the estimation of the expeced layer loss $\mathrm{E}(\hat{S}_q)$ in the prospective treaty year q before AAD/AAL.

DAV Deutsche Aktuarvereinigung e.V. DGVFM Deutsche Gesellschaft für Versicherungs- und Finanzmathematik e.V.

Burning Cost (BC)

An important step in the pricing of the layer C xs D is the estimation of the expeced layer loss $\mathrm{E}(\hat{S}_q)$ in the prospective treaty year q before AAD/AAL.

Burning cost (BC) estimator:

$$\widehat{\mathrm{E}}^{\mathrm{BC}}_{C \text{ xs } D} := v_q \cdot \frac{\sum_{i=1}^{n} \sum_{k=1}^{N_i} \min(C, (X_{i,k} - D)^+)}{\sum_{i=1}^{n} v_n}$$

Burning Cost (BC)

An important step in the pricing of the layer C xs D is the estimation of the expeced layer loss $\mathrm{E}(\hat{S}_q)$ in the prospective treaty year q before AAD/AAL.

Burning cost (BC) estimator:

$$\widehat{\mathrm{E}}^{\mathrm{BC}}_{C \text{ xs } D} := v_q \cdot \frac{\sum_{i=1}^{n} \sum_{k=1}^{N_i} \min(C, (X_{i,k} - D)^+)}{\sum_{i=1}^{n} v_n}$$

Assumptions:

- The accident years are independent

Burning Cost (BC)

An important step in the pricing of the layer C xs D is the estimation of the expeced layer loss $\mathrm{E}(\hat{S}_q)$ in the prospective treaty year q before AAD/AAL.

Burning cost (BC) estimator:

$$\widehat{\mathrm{E}}^{\mathrm{BC}}_{C \text{ xs } D} := v_q \cdot \frac{\sum_{i=1}^{n} \sum_{k=1}^{N_i} \min(C, (X_{i,k} - D)^+)}{\sum_{i=1}^{n} v_n}$$

Assumptions:

- The accident years are independent
- The random sums $\sum_{k=1}^{N_i} X_{i,k}$ are collective models for all accident years i

Burning Cost (BC)

An important step in the pricing of the layer C xs D is the estimation of the expeced layer loss $\mathrm{E}(\hat{S}_q)$ in the prospective treaty year q before AAD/AAL.

Burning cost (BC) estimator:

$$\widehat{\mathrm{E}}^{\mathrm{BC}}_{C \text{ xs } D} := v_q \cdot \frac{\sum_{i=1}^{n} \sum_{k=1}^{N_i} \min(C, (X_{i,k} - D)^+)}{\sum_{i=1}^{n} v_n}$$

Assumptions:

- The accident years are independent
- The random sums $\sum_{k=1}^{N_i} X_{i,k}$ are collective models for all accident years i
- There is a $\lambda > 0$, such that $N_i \sim \mathrm{Poisson}(\lambda v_i)$ for all i

Burning Cost (BC)

An important step in the pricing of the layer C xs D is the estimation of the expeced layer loss $\mathrm{E}(\hat{S}_q)$ in the prospective treaty year q before AAD/AAL.

Burning cost (BC) estimator:

$$\widehat{\mathrm{E}}^{\mathrm{BC}}_{C\ \mathrm{xs}\ D} := v_q \cdot \frac{\sum_{i=1}^{n} \sum_{k=1}^{N_i} \min(C, (X_{i,k} - D)^+)}{\sum_{i=1}^{n} v_n}$$

Assumptions:

- The accident years are independent
- The random sums $\sum_{k=1}^{N_i} X_{i,k}$ are collective models for all accident years i
- There is a $\lambda > 0$, such that $N_i \sim \mathrm{Poisson}(\lambda v_i)$ for all i
- The loss sizes $X_{i,k}$ are identically distributed for all accident years i

Burning Cost (BC)

An important step in the pricing of the layer C xs D is the estimation of the expeced layer loss $\mathrm{E}(\hat{S}_q)$ in the prospective treaty year q before AAD/AAL.

Burning cost (BC) estimator:

$$\widehat{\mathrm{E}}^{\mathrm{BC}}_{C \text{ xs } D} := v_q \cdot \frac{\sum_{i=1}^{n} \sum_{k=1}^{N_i} \min(C, (X_{i,k} - D)^+)}{\sum_{i=1}^{n} v_n}$$

Assumptions:

- The accident years are independent
- The random sums $\sum_{k=1}^{N_i} X_{i,k}$ are collective models for all accident years i
- There is a $\lambda > 0$, such that $N_i \sim \text{Poisson}(\lambda v_i)$ for all i
- The loss sizes $X_{i,k}$ are identically distributed for all accident years i

Under these assumptions, the BC estimator is unbiased! The assumptions are plausible if indexation works perfectly and if the v_i are good volumes.

Agenda

1. Burning cost for per risk XLs
2. Pareto based pricing approaches
3. Simulation model for evaluating pricing approaches
4. Observations with underlying Pareto severity
5. Observations with other underlying severity distributions

Pareto distribution

Die Pareto distribution Pareto(t, α) is the most widely used severity distribution for modelling large losses. CDF:

$$F(x) = 1 - \left(\frac{t}{x}\right)^{\alpha}, \quad x \geq t.$$

Proposition:

- Let $X \sim \text{Pareto}(t, \alpha)$. Then $cX \sim \text{Pareto}(ct, \alpha)$ for all $c > 0$.

Pareto distribution

Die Pareto distribution Pareto(t, α) is the most widely used severity distribution for modelling large losses. CDF:

$$F(x) = 1 - \left(\frac{t}{x}\right)^{\alpha}, \quad x \geq t.$$

Proposition:

- Let $X \sim$ Pareto(t, α). Then $cX \sim$ Pareto(ct, α) for all $c > 0$.
- Let $X \sim$ Pareto(t_1, α). For $t_2 > t_1$ we then have $X|(X > t_2) \sim$ Pareto(t_2, α).

Pareto distribution

Die Pareto distribution Pareto(t, α) is the most widely used severity distribution for modelling large losses. CDF:

$$F(x) = 1 - \left(\frac{t}{x}\right)^{\alpha}, \quad x \geq t.$$

Proposition:

- Let $X \sim$ Pareto(t, α). Then $cX \sim$ Pareto(ct, α) for all $c > 0$.
- Let $X \sim$ Pareto(t_1, α). For $t_2 > t_1$ we then have $X|(X > t_2) \sim$ Pareto(t_2, α).

Consequences:

- The Pareto alpha is invariant wrt scaling (in particular wrt FX conversion and inflation)

Pareto distribution

Die Pareto distribution Pareto(t, α) is the most widely used severity distribution for modelling large losses. CDF:

$$F(x) = 1 - \left(\frac{t}{x}\right)^{\alpha}, \quad x \geq t.$$

Proposition:

- Let $X \sim$ Pareto(t, α). Then $cX \sim$ Pareto(ct, α) for all $c > 0$.
- Let $X \sim$ Pareto(t_1, α). For $t_2 > t_1$ we then have $X|(X > t_2) \sim$ Pareto(t_2, α).

Consequences:

- The Pareto alpha is invariant wrt scaling (in particular wrt FX conversion and inflation)
- For Pareto distributed data, the Pareto alpha does not depend on the reporting threshold

Pareto distribution

Die Pareto distribution Pareto(t, α) is the most widely used severity distribution for modelling large losses. CDF:

$$F(x) = 1 - \left(\frac{t}{x}\right)^{\alpha}, \quad x \geq t.$$

Proposition:

- Let $X \sim$ Pareto(t, α). Then $cX \sim$ Pareto(ct, α) for all $c > 0$.
- Let $X \sim$ Pareto(t_1, α). For $t_2 > t_1$ we then have $X|(X > t_2) \sim$ Pareto(t_2, α).

Consequences:

- The Pareto alpha is invariant wrt scaling (in particular wrt FX conversion and inflation)
- For Pareto distributed data, the Pareto alpha does not depend on the reporting threshold
- The Pareto distribution has only one ‘real’ parameter: the Pareto alpha

Pareto distribution

Die Pareto distribution Pareto(t, α) is the most widely used severity distribution for modelling large losses. CDF:

$$F(x) = 1 - \left(\frac{t}{x}\right)^{\alpha}, \quad x \geq t.$$

Proposition:

- Let $X \sim$ Pareto(t, α). Then $cX \sim$ Pareto(ct, α) for all $c > 0$.
- Let $X \sim$ Pareto(t_1, α). For $t_2 > t_1$ we then have $X|(X > t_2) \sim$ Pareto(t_2, α).

Consequences:

- The Pareto alpha is invariant wrt scaling (in particular wrt FX conversion and inflation)
- For Pareto distributed data, the Pareto alpha does not depend on the reporting threshold
- The Pareto distribution has only one 'real' parameter: the Pareto alpha

Due to these properties, the Pareto distribution is very useful for comparing tails. It is often possible to provide 'market alphas' for various segments.

Pareto distribution

Proposition (Maximum likelihood estimation of α):

Let $t > 0$ and $X_1, \ldots, X_n \sim \text{Pareto}(t, \alpha)$. The *maximum likelihood estimator (MLE)* for α is given by

$$\widehat{\alpha}^{\text{ML}} := \frac{n}{\sum_{i=1}^{n} \ln(X_i/t)}.$$

Pareto distribution

Proposition (Maximum likelihood estimation of α):

Let $t > 0$ and $X_1, \dots, X_n \sim \text{Pareto}(t, \alpha)$. The *maximum likelihood estimator (MLE)* for α is given by

$$\widehat{\alpha}^{\text{ML}} := \frac{n}{\sum_{i=1}^{n} \ln(X_i/t)}.$$

Alternatively, one can use the *bias corrected MLE*

$$\widehat{\alpha}^{*} := \frac{n-1}{n} \cdot \widehat{\alpha}^{\text{ML}} = \frac{n-1}{\sum_{i=1}^{n} \ln(X_i/t)}.$$

Pareto distribution

Proposition (Maximum likelihood estimation of α):

Let $t > 0$ and $X_1, \ldots, X_n \sim \text{Pareto}(t, \alpha)$. The *maximum likelihood estimator (MLE)* for α is given by

$$\widehat{\alpha}^{\text{ML}} := \frac{n}{\sum_{i=1}^{n} \ln(X_i/t)}.$$

Alternatively, one can use the *bias corrected MLE*

$$\widehat{\alpha}^{*} := \frac{n-1}{n} \cdot \widehat{\alpha}^{\text{ML}} = \frac{n-1}{\sum_{i=1}^{n} \ln(X_i/t)}.$$

The bias corrected MLE $\widehat{\alpha}^{*}$ is unbiased and has a lower variance than $\widehat{\alpha}^{\text{ML}}$. Hence, $\widehat{\alpha}^{*}$ is a better estimator than $\widehat{\alpha}^{\text{ML}}$.

M. Rytgaard (1990) Estimation in the Pareto Distribution, ASTIN Bullletin: The Journal of the IAA, 20(2), pp. 201–216

Pareto distribution

Proposition (layer mean): Let $X \sim \text{Pareto}(t, \alpha)$. For a layer C xs D with $D \geq t$ we then have:

$$\text{E}(\min(C; (X-D)^+)) = \begin{cases} \frac{t^\alpha}{1-\alpha}\left((C+D)^{1-\alpha} - D^{1-\alpha}\right) & \text{if } \alpha \neq 1 \\ t\left(\ln(C+D) - \ln(D)\right) & \text{if } \alpha = 1. \end{cases}$$

Pareto distribution

Proposition (layer mean): Let $X \sim \text{Pareto}(t, \alpha)$. For a layer C xs D with $D \geq t$ we then have:

$$\mathrm{E}(\min(C; (X - D)^+)) = \begin{cases} \frac{t^\alpha}{1-\alpha}\left((C+D)^{1-\alpha} - D^{1-\alpha}\right) & \text{if } \alpha \neq 1 \\ t\left(\ln(C+D) - \ln(D)\right) & \text{if } \alpha = 1. \end{cases}$$

Proposition (Pareto extrapolation): Let $\sum_{n=1}^{N} X_n$ be a collective model with $X_n \sim X \sim \text{Pareto}(t, \alpha)$. For layers C_i xs D_i with $D_i \geq t$ we then have

$$\Phi_{C_1 \text{ xs } D_1}^{C_2 \text{ xs } D_2}(\alpha) := \frac{\text{Exp. loss in } C_2 \text{ xs } D_2}{\text{Exp. loss in } C_1 \text{ xs } D_1} = \begin{cases} \frac{(C_2 + D_2)^{1-\alpha} - D_2^{1-\alpha}}{(C_1 + D_1)^{1-\alpha} - D_1^{1-\alpha}} & \text{if } \alpha \neq 1 \\ \frac{\ln(C_2 + D_2) - \ln(D_2)}{\ln(C_1 + D_1) - \ln(D_1)} & \text{if } \alpha = 1 \end{cases}$$

Approach 1: Poisson/Pareto modell

For pricing C xs D select a large loss threshold $t \leq D$ and model the large losses in the prospective treaty year q with a collective model $\sum_{i=1}^{N_q} X_{q,i}$:

Approach 1: Poisson/Pareto modell

For pricing C xs D select a large loss threshold $t \leq D$ and model the large losses in the prospective treaty year q with a collective model $\sum_{i=1}^{N_q} X_{q,i}$:

- $X_{q,i} \sim \text{Pareto}(t, \widehat{\alpha})$, where $\widehat{\alpha}$ denotes the MLE or the bias corrected MLE for α (calculated with the indexed large losses from the observation period $\{1, \ldots, n\}$).

Approach 1: Poisson/Pareto modell

For pricing C xs D select a large loss threshold $t \leq D$ and model the large losses in the prospective treaty year q with a collective model $\sum_{i=1}^{N_q} X_{q,i}$:

- $X_{q,i} \sim \text{Pareto}(t, \widehat{\alpha})$, where $\widehat{\alpha}$ denotes the MLE or the bias corrected MLE for α (calculated with the indexed large losses from the observation period $\{1, \ldots, n\}$).
- $N_q \sim \text{Poisson}(\widehat{\lambda}\, v_q)$, where

$$\widehat{\lambda} := \frac{\sum_{i=1}^{n} \sum_{k=1}^{N_i} \chi_{\{X_{i,k} > t\}}}{\sum_{i=1}^{n} v_i}$$

denotes the MLE of λ.

Approach 1: Poisson/Pareto modell

For pricing C xs D select a large loss threshold $t \leq D$ and model the large losses in the prospective treaty year q with a collective model $\sum_{i=1}^{N_q} X_{q,i}$:

- $X_{q,i} \sim \text{Pareto}(t, \widehat{\alpha})$, where $\widehat{\alpha}$ denotes the MLE or the bias corrected MLE for α (calculated with the indexed large losses from the observation period $\{1, \ldots, n\}$).
- $N_q \sim \text{Poisson}(\widehat{\lambda}\, v_q)$, where

$$\widehat{\lambda} := \frac{\sum_{i=1}^{n} \sum_{k=1}^{N_i} \chi_{\{X_{i,k} > t\}}}{\sum_{i=1}^{n} v_i}$$

  denotes the MLE of λ.

The expected loss in C xs D (before AAD/AAL) is then estimated by

$$\widehat{\mathrm{E}}_{C \text{ xs } D}^{\mathrm{PP}} := \widehat{\lambda}\, v_q \cdot \mathrm{E}(\min(C; (X_{q,1} - D)^{+})).$$

Approach 2: Pareto extrapolation

Select a large loss threshold t like in Approach 1 and a $T > t$, such that the BC estimator $\widehat{\mathrm{E}}^{\mathrm{BC}}_{c_0 \text{ xs } d_0}$ is reasonable for the layer c_0 xs d_0 with $d_0 := t$ and $c_0 := T - t$.

Approach 2: Pareto extrapolation

Select a large loss threshold t like in Approach 1 and a $T > t$, such that the BC estimator $\widehat{\mathrm{E}}^{\mathrm{BC}}_{c_0 \text{ xs } d_0}$ is reasonable for the layer c_0 xs d_0 with $d_0 := t$ and $c_0 := T - t$. Choose an estimator $\widehat{\alpha}$ for α.

Approach 2: Pareto extrapolation

Select a large loss threshold t like in Approach 1 and a $T > t$, such that the BC estimator $\widehat{\mathrm{E}}^{\mathrm{BC}}_{c_0 \text{ xs } d_0}$ is reasonable for the layer c_0 xs d_0 with $d_0 := t$ and $c_0 := T - t$. Choose an estimator $\widehat{\alpha}$ for α.

Estimator for the expected layer loss in C xs D (befor AAD/AAL):

- Case $C + D \leq T$: Use the BC estimator

$$\widehat{\mathrm{E}}^{\mathrm{PE}}_{C \text{ xs } D} := \widehat{\mathrm{E}}^{\mathrm{BC}}_{C \text{ xs } D}$$

Approach 2: Pareto extrapolation

Select a large loss threshold t like in Approach 1 and a $T > t$, such that the BC estimator $\widehat{\mathrm{E}}^{\mathrm{BC}}_{c_0 \text{ xs } d_0}$ is reasonable for the layer c_0 xs d_0 with $d_0 := t$ and $c_0 := T - t$. Choose an estimator $\widehat{\alpha}$ for α.

Estimator for the expected layer loss in C xs D (befor AAD/AAL):

- Case $C + D \leq T$: Use the BC estimator

$$\widehat{\mathrm{E}}^{\mathrm{PE}}_{C \text{ xs } D} := \widehat{\mathrm{E}}^{\mathrm{BC}}_{C \text{ xs } D}$$

- Case $T \leq D$: Use Pareto extrapolation from the layer c_0 xs d_0 for the entire layer

$$\widehat{\mathrm{E}}^{\mathrm{PE}}_{C \text{ xs } D} := \widehat{\mathrm{E}}^{\mathrm{BC}}_{c_0 \text{ xs } d_0} \cdot \Phi^{C \text{ xs } D}_{c_0 \text{ xs } d_0}(\widehat{\alpha})$$

Approach 2: Pareto extrapolation

Select a large loss threshold t like in Approach 1 and a $T > t$, such that the BC estimator $\widehat{\mathrm{E}}^{\mathrm{BC}}_{c_0 \text{ xs } d_0}$ is reasonable for the layer c_0 xs d_0 with $d_0 := t$ and $c_0 := T - t$. Choose an estimator $\widehat{\alpha}$ for α.

Estimator for the expected layer loss in C xs D (befor AAD/AAL):

- Case $C + D \leq T$: Use the BC estimator

$$\widehat{\mathrm{E}}^{\mathrm{PE}}_{C \text{ xs } D} := \widehat{\mathrm{E}}^{\mathrm{BC}}_{C \text{ xs } D}$$

- Case $T \leq D$: Use Pareto extrapolation from the layer c_0 xs d_0 for the entire layer

$$\widehat{\mathrm{E}}^{\mathrm{PE}}_{C \text{ xs } D} := \widehat{\mathrm{E}}^{\mathrm{BC}}_{c_0 \text{ xs } d_0} \cdot \Phi^{C \text{ xs } D}_{c_0 \text{ xs } d_0}(\widehat{\alpha})$$

- Case $T \in (D, C + D)$: Use Pareto extrapolation for the layer part above T and BC for the layer part below T

$$\widehat{\mathrm{E}}^{\mathrm{PE}}_{C \text{ xs } D} := \widehat{\mathrm{E}}^{\mathrm{BC}}_{T-D \text{ xs } D} + \widehat{\mathrm{E}}^{\mathrm{BC}}_{c_0 \text{ xs } d_0} \cdot \Phi^{C+D-T \text{ xs } T}_{c_0 \text{ xs } d_0}(\widehat{\alpha})$$

Variants of the Approaches 1 and 2

Variant 1: Assume there is market knowledge about a ‘typical’ alpha in the considered segment. Then, the pricig actuary will not use fitted alphas that deviate too much from the typical alpha.

Variants of the Approaches 1 and 2

Variant 1: Assume there is market knowledge about a ‘typical’ alpha in the considered segment. Then, the pricig actuary will not use fitted alphas that deviate too much from the typical alpha. This corresponds to the selection of $\alpha_{\text{min}} \leq \alpha_{\text{max}}$ and the use of

$$\min(\alpha_{\text{max}}, \max(\alpha_{\text{min}}, \widehat{\alpha}^{\text{ML}})) \quad \text{or} \quad \min(\alpha_{\text{max}}, \max(\alpha_{\text{min}}, \widehat{\alpha}^{*})).$$

The extreme case $\alpha_{\text{min}} = \alpha_{\text{max}}$ corresponds to the use of a market alpha.

Variants of the Approaches 1 and 2

Variant 1: Assume there is market knowledge about a 'typical' alpha in the considered segment. Then, the pricig actuary will not use fitted alphas that deviate too much from the typical alpha. This corresponds to the selection of $\alpha_{\text{min}} \leq \alpha_{\text{max}}$ and the use of

$$\min(\alpha_{\text{max}}, \max(\alpha_{\text{min}}, \widehat{\alpha}^{\text{ML}})) \quad \text{or} \quad \min(\alpha_{\text{max}}, \max(\alpha_{\text{min}}, \widehat{\alpha}^{*})).$$

The extreme case $\alpha_{\text{min}} = \alpha_{\text{max}}$ corresponds to the use of a market alpha.

Variant 2 (BC/Pareto): Like in Aproach 2, select a $T > t$ (depending on the loss history) and use the BC estimator below T and the Poisson/Pareto modell from Approach 1 above T.

Variants of the Approaches 1 and 2

Variant 1: Assume there is market knowledge about a 'typical' alpha in the considered segment. Then, the pricig actuary will not use fitted alphas that deviate too much from the typical alpha. This corresponds to the selection of $\alpha_{\text{min}} \leq \alpha_{\text{max}}$ and the use of

$$\min(\alpha_{\text{max}}, \max(\alpha_{\text{min}}, \widehat{\alpha}^{\text{ML}})) \quad \text{or} \quad \min(\alpha_{\text{max}}, \max(\alpha_{\text{min}}, \widehat{\alpha}^{*})).$$

The extreme case $\alpha_{\text{min}} = \alpha_{\text{max}}$ corresponds to the use of a market alpha.

Variant 2 (BC/Pareto): Like in Aproach 2, select a $T > t$ (depending on the loss history) and use the BC estimator below T and the Poisson/Pareto modell from Approach 1 above T.

Rules of thumb for the selection of T:

- T = largest loss
- T = 3rd larges loss
- $T = \ldots$

Agenda

Simulation model for evaluating pricing approaches

Choose underlying model (‘reality’)

Simulation model for evaluating pricing approaches

Choose underlying model (‘reality’)

↓

Simulate (for instance) 20,000 observation periods with the underlying model

Simulation model for evaluating pricing approaches

Choose underlying model ('reality')

↓

Simulate (for instance) 20,000 observation periods with the underlying model

↓

Apply the pricing approaches for each of the simulated observation periods

Simulation model for evaluating pricing approaches

Choose underlying model (‘reality’)

↓

Simulate (for instance) 20,000 observation periods with the underlying model

↓

Apply the pricing approaches for each of the simulated observation periods

↓

Calculate the squared deviation between the expected layer loss of the underlying model and the estimators of the pricing models

Simulation model for evaluating pricing approaches

Choose underlying model (‘reality’)

↓

Simulate (for instance) 20,000 observation periods with the underlying model

↓

Apply the pricing approaches for each of the simulated observation periods

↓

Calculate the squared deviation between the expected layer loss of the underlying model and the estimators of the pricing models

↓

Calculate the square roots of the average squared deviations to estimate the standard errors (RMSE) of the pricing methods

Agenda

1. Burning cost for per risk XLs
2. Pareto based pricing approaches
3. Simulation model for evaluating pricing approaches
4. Observations with underlying Pareto severity
5. Observations with other underlying severity distributions

Simulation settings

For simplicity, we use the assumption

$$v_q = \sum_{i=1}^{n} v_i = 1.$$

Simulation settings

For simplicity, we use the assumption

$$v_q = \sum_{i=1}^{n} v_i = 1.$$

We simulate observation periods with

- claim count $N \sim \text{Poisson}(\lambda) \quad (\lambda \in \{5, 15, 50\})$

Simulation settings

For simplicity, we use the assumption

$$v_q = \sum_{i=1}^{n} v_i = 1.$$

We simulate observation periods with

- claim count $N \sim \text{Poisson}(\lambda) \quad (\lambda \in \{5, 15, 50\})$
- and severity $X_n \sim \text{Pareto}(t, \alpha)$ with $t = 1{,}000$ and $\alpha = 2.0$

Simulation settings

For simplicity, we use the assumption

$$v_q = \sum_{i=1}^{n} v_i = 1.$$

We simulate observation periods with

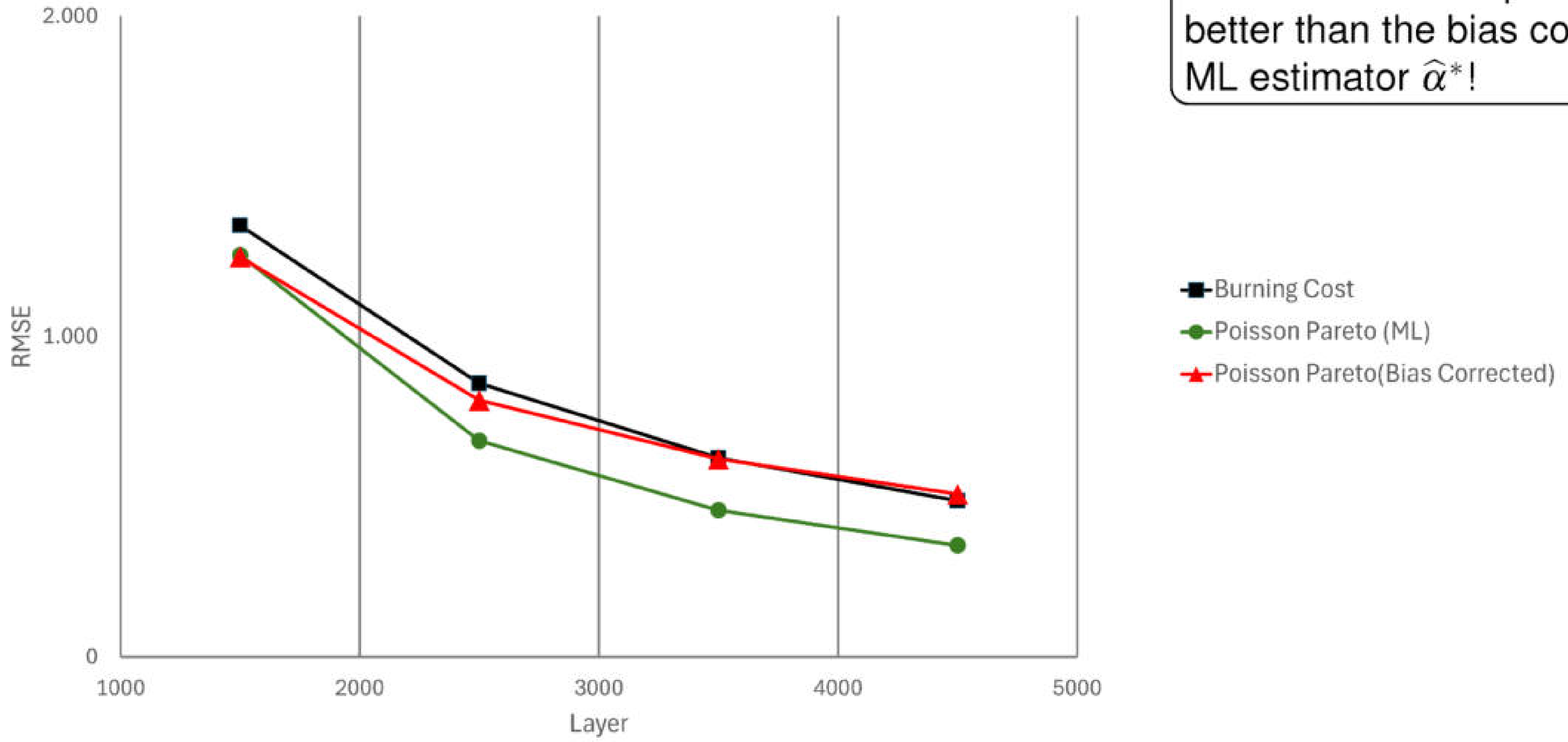
- claim count $N \sim \text{Poisson}(\lambda) \quad (\lambda \in \{5, 15, 50\})$
- and severity $X_n \sim \text{Pareto}(t, \alpha)$ with $t = 1{,}000$ and $\alpha = 2.0$

and use the pricing approaches to estimate the expected losses of the following layers:

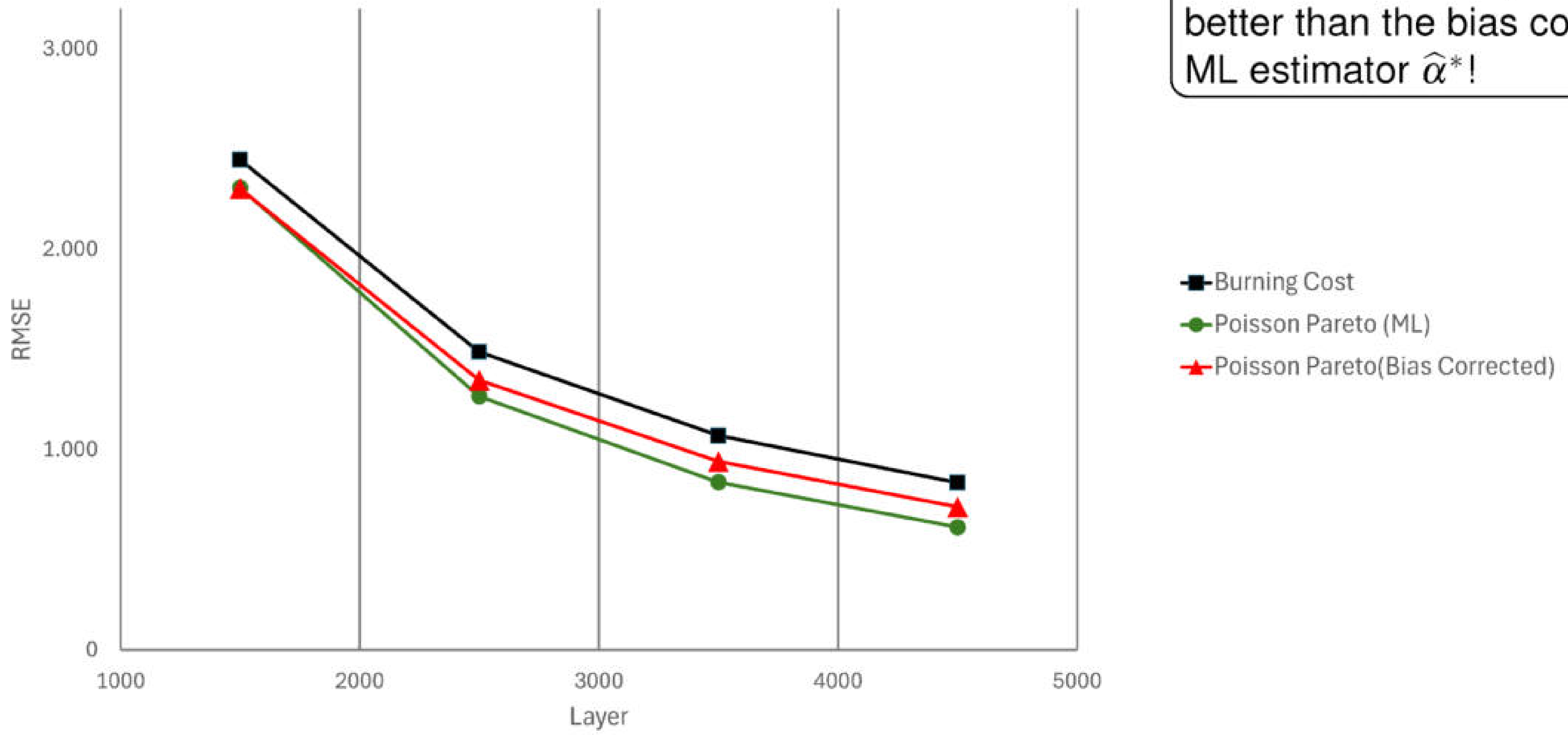
- 1,000 xs 1,000
- 1,000 xs 2,000
- 1,000 xs 3,000
- 1,000 xs 4,000

RMSE with $\lambda = 5$, $\alpha = 2.0$

ML estimator $\widehat{\alpha}^{ML}$ performs better than the bias corrected ML estimator $\widehat{\alpha}^{*}$!



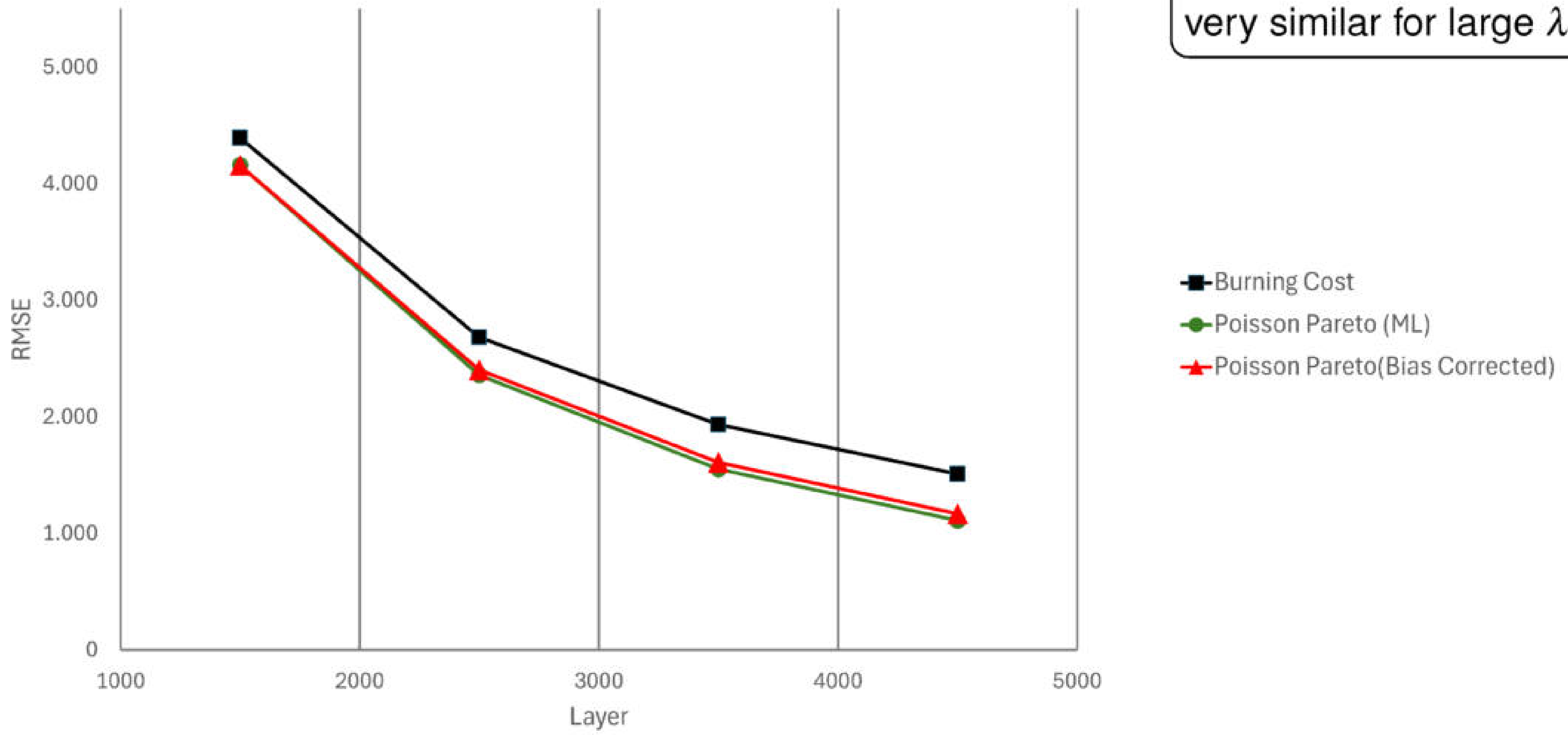
RMSE with $\lambda = 15$, $\alpha = 2.0$



ML estimator $\widehat{\alpha}^{\mathrm{ML}}$ performs better than the bias corrected ML estimator $\widehat{\alpha}^{*}$!

RMSE with $\lambda = 50$, $\alpha = 2.0$

The estimators $\widehat{\alpha}^{\mathrm{ML}}$ and $\widehat{\alpha}^{*}$ are very similar for large λ.



Observation 1

The Pareto based methods seem to perform better than the Burning Cost.

Observation 1

The Pareto based methods seem to perform better than the Burning Cost.

This is not surprising, as the underlying distribution is Pareto . . .

Observation 2

The ML estimator $\widehat{\alpha}^{\mathrm{ML}}$ typically performs better than the bias corrected ML estimator $\widehat{\alpha}^{*}$.

Observation 2

The ML estimator $\widehat{\alpha}^{\mathrm{ML}}$ typically performs better than the bias corrected ML estimator $\widehat{\alpha}^{*}$.

At first sight, this is surprising as $\widehat{\alpha}^{*}$ a the better estimator for α than $\widehat{\alpha}^{\mathrm{ML}}$!

Observation 2

The ML estimator $\widehat{\alpha}^{\mathrm{ML}}$ typically performs better than the bias corrected ML estimator $\widehat{\alpha}^*$.

At first sight, this is surprising as $\widehat{\alpha}^*$ a the better estimator for α than $\widehat{\alpha}^{\mathrm{ML}}$!

Explanation:

Let $X_\alpha \sim \mathrm{Pareto}(t, \alpha)$ with $t \le D$. Then the function

$$\alpha \mapsto \mu(\alpha) := \mathrm{E}(\min(C, (X_\alpha - D)^+)$$

is convex!

Observation 2

The ML estimator $\widehat{\alpha}^{\mathrm{ML}}$ typically performs better than the bias corrected ML estimator $\widehat{\alpha}^*$.

At first sight, this is surprising as $\widehat{\alpha}^*$ a the better estimator for α than $\widehat{\alpha}^{\mathrm{ML}}$!

Explanation:

Let $X_\alpha \sim \mathrm{Pareto}(t, \alpha)$ with $t \leq D$. Then the function

$$\alpha \mapsto \mu(\alpha) := \mathrm{E}(\min(C, (X_\alpha - D)^+)$$

is convex! Jensen's inequality implies

$$\mathrm{E}(\mu(\widehat{\alpha}^*)) > \mu(\mathrm{E}(\widehat{\alpha}^*)).$$

Observation 2

The ML estimator $\widehat{\alpha}^{\mathrm{ML}}$ typically performs better than the bias corrected ML estimator $\widehat{\alpha}^*$.

At first sight, this is surprising as $\widehat{\alpha}^*$ a the better estimator for α than $\widehat{\alpha}^{\mathrm{ML}}$!

Explanation:

Let $X_\alpha \sim \mathrm{Pareto}(t, \alpha)$ with $t \leq D$. Then the function

$$\alpha \mapsto \mu(\alpha) := \mathrm{E}(\min(C, (X_\alpha - D)^+)$$

is convex! Jensen's inequality implies

$$\mathrm{E}(\mu(\widehat{\alpha}^*)) > \mu(\mathrm{E}(\widehat{\alpha}^*)).$$

$\Rightarrow$ Estimation of the expected layer loss with $\widehat{\alpha}^*$ has a positive bias!

DAV Deutsche Aktuarvereinigung e.V. DGVFM Deutsche Gesellschaft für Versicherungs- und Finanzmathematik e.V.

Observation 2

The ML estimator $\widehat{\alpha}^{\mathrm{ML}}$ typically performs better than the bias corrected ML estimator $\widehat{\alpha}^*$.

At first sight, this is surprising as $\widehat{\alpha}^*$ a the better estimator for α than $\widehat{\alpha}^{\mathrm{ML}}$!

Explanation:

Let $X_\alpha \sim \mathrm{Pareto}(t, \alpha)$ with $t \leq D$. Then the function

$$\alpha \mapsto \mu(\alpha) := \mathrm{E}(\min(C, (X_\alpha - D)^+)$$

is convex! Jensen's inequality implies

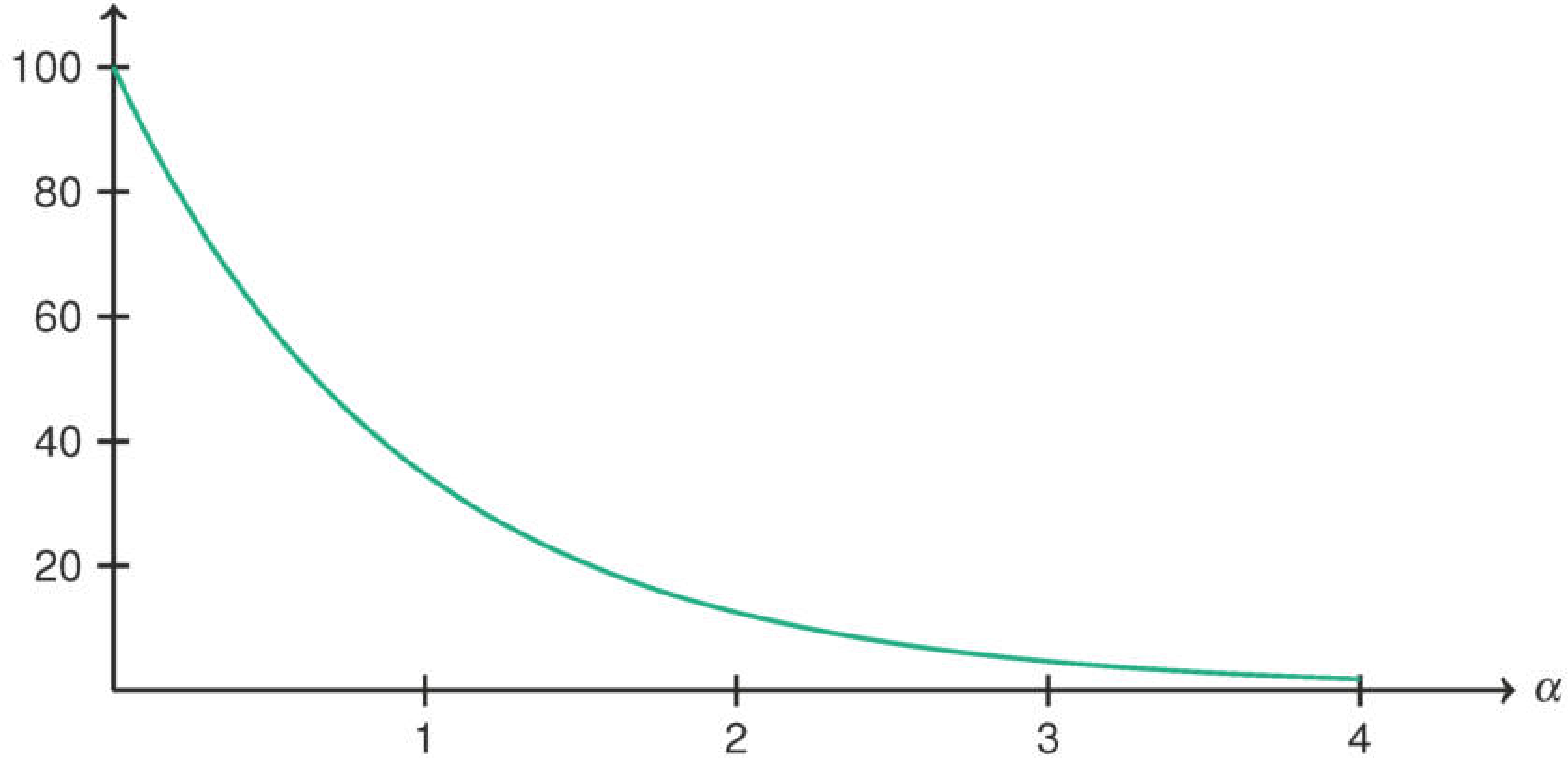
$$\mathrm{E}(\mu(\widehat{\alpha}^*)) > \mu(\mathrm{E}(\widehat{\alpha}^*)).$$

$\Rightarrow$ Estimation of the expected layer loss with $\widehat{\alpha}^*$ has a positive bias!

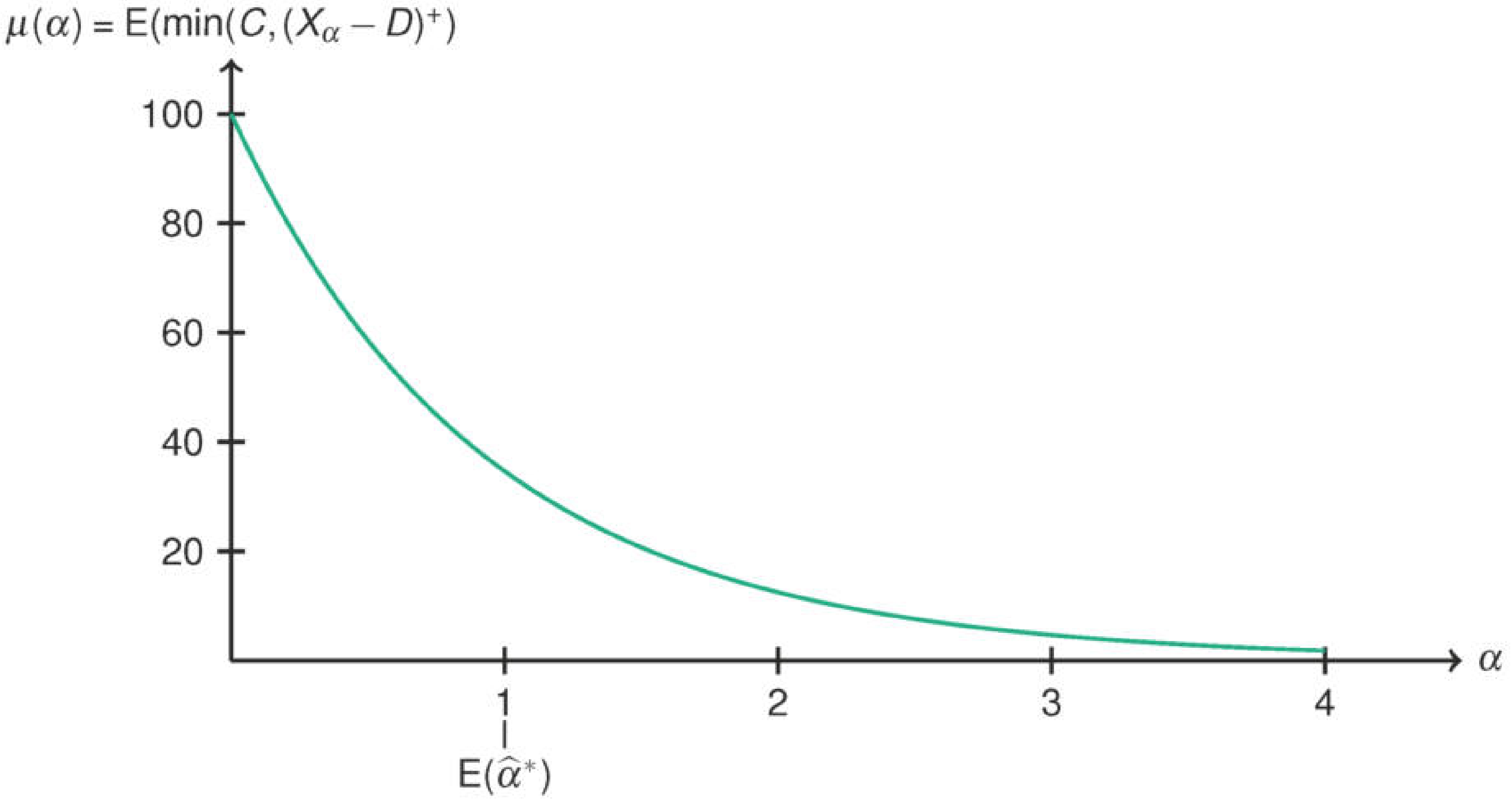
The ML estimator $\widehat{\alpha}^{\mathrm{ML}}$ results in lower estimates for the expected layer loss, which compensates this positive bias to some extent!

Jensen's Inequality

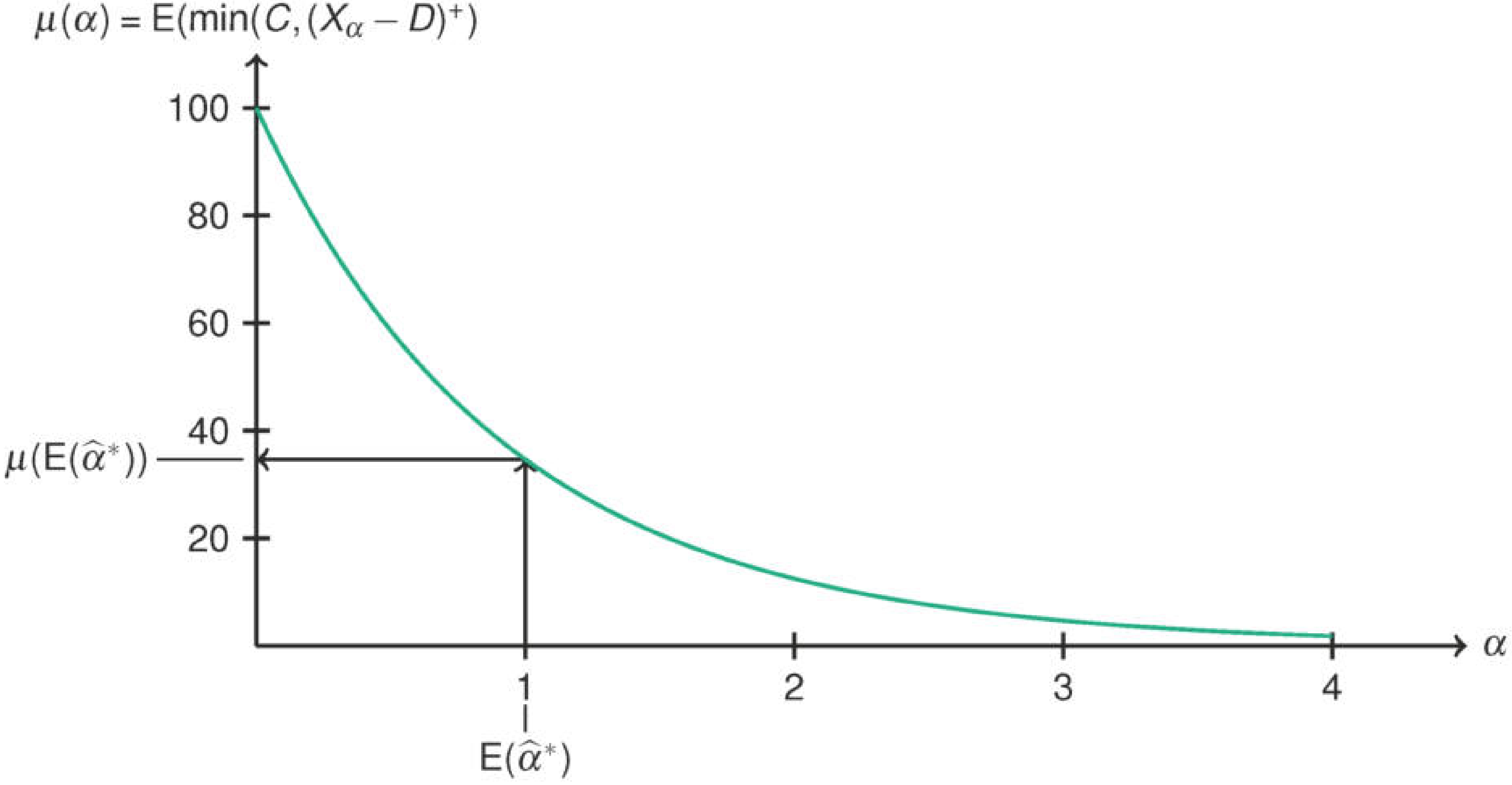
$\mu(\alpha) = \mathrm{E}(\min(C, (X_\alpha - D)^+)$



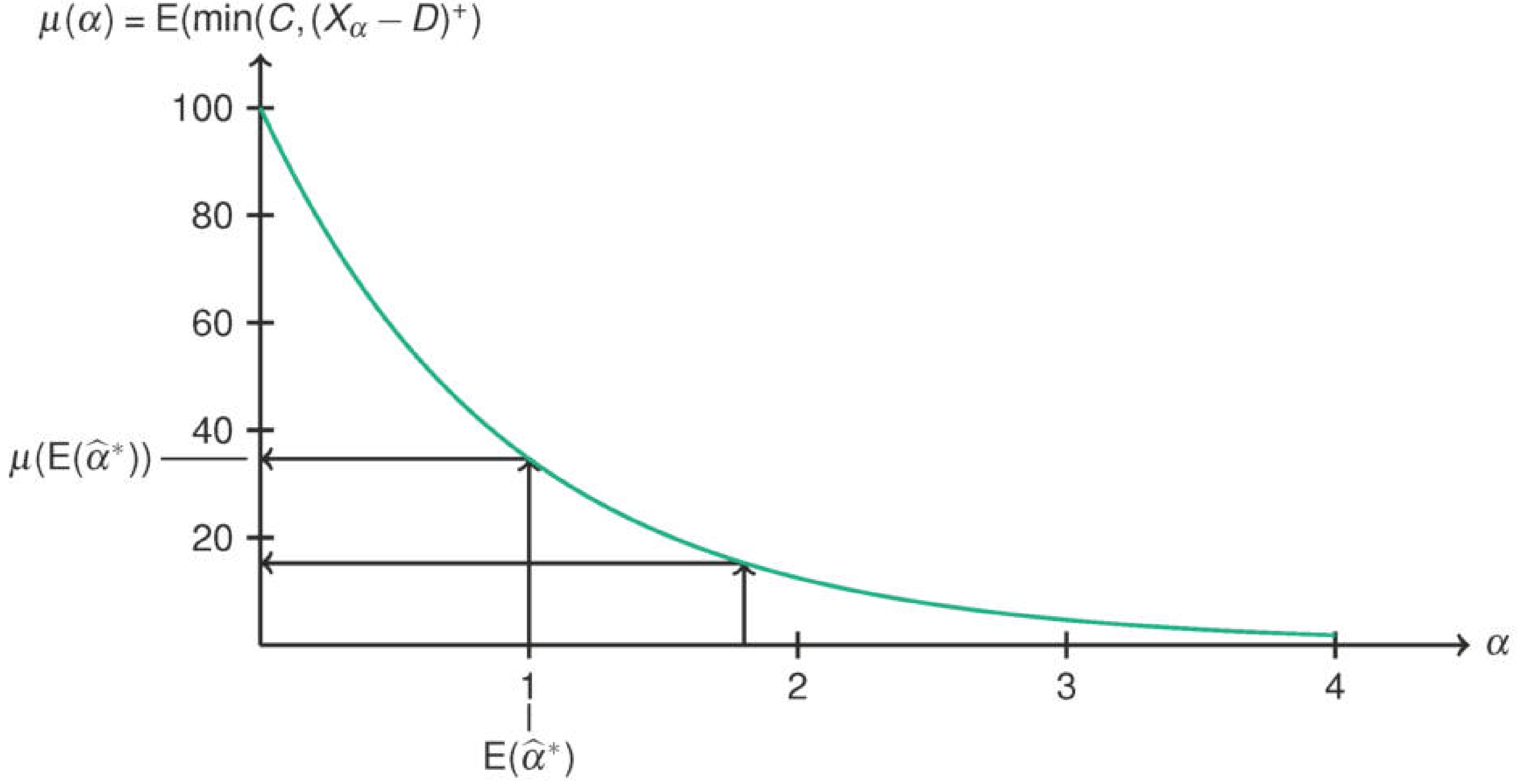
Jensen's Inequality



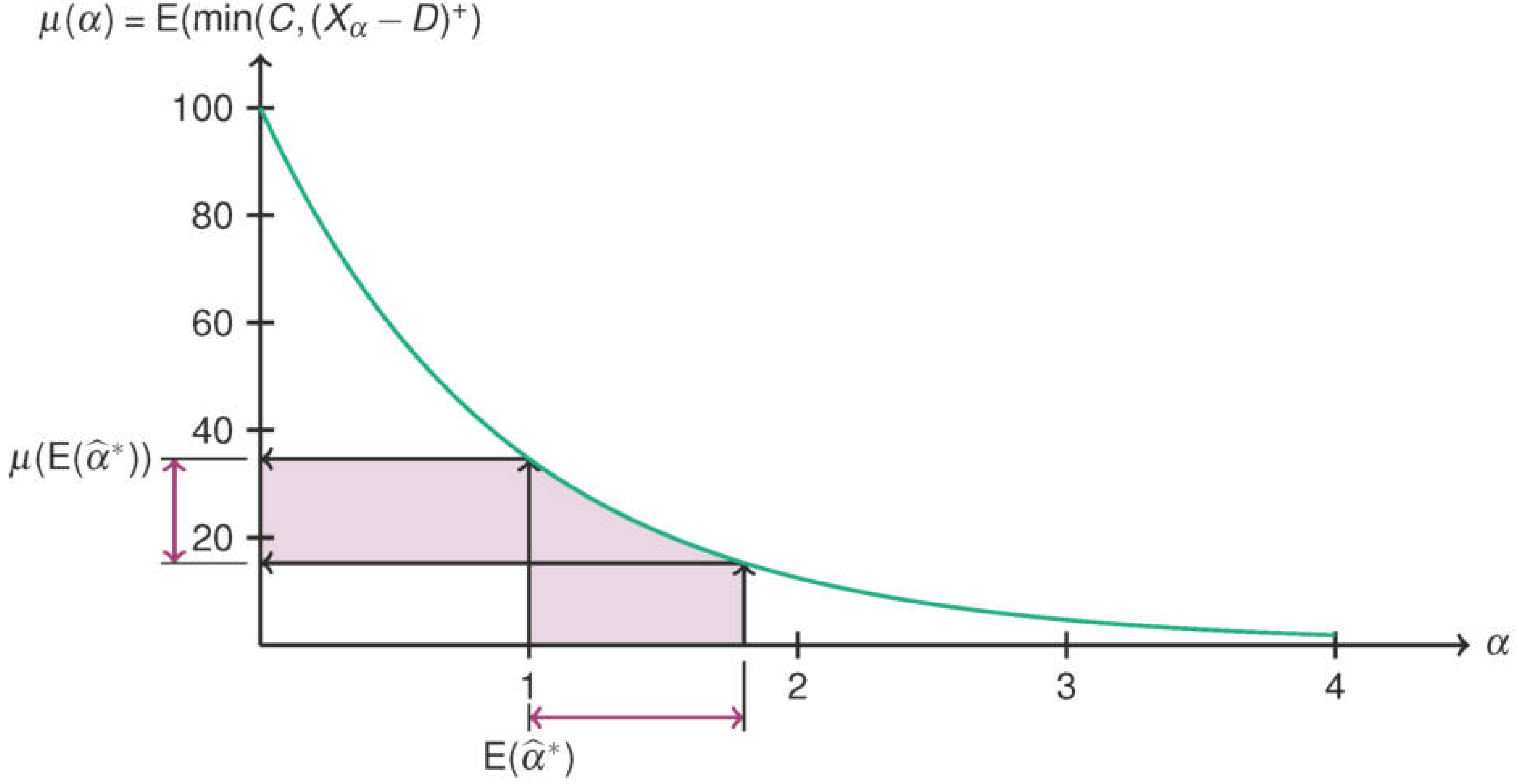
Jensen's Inequality



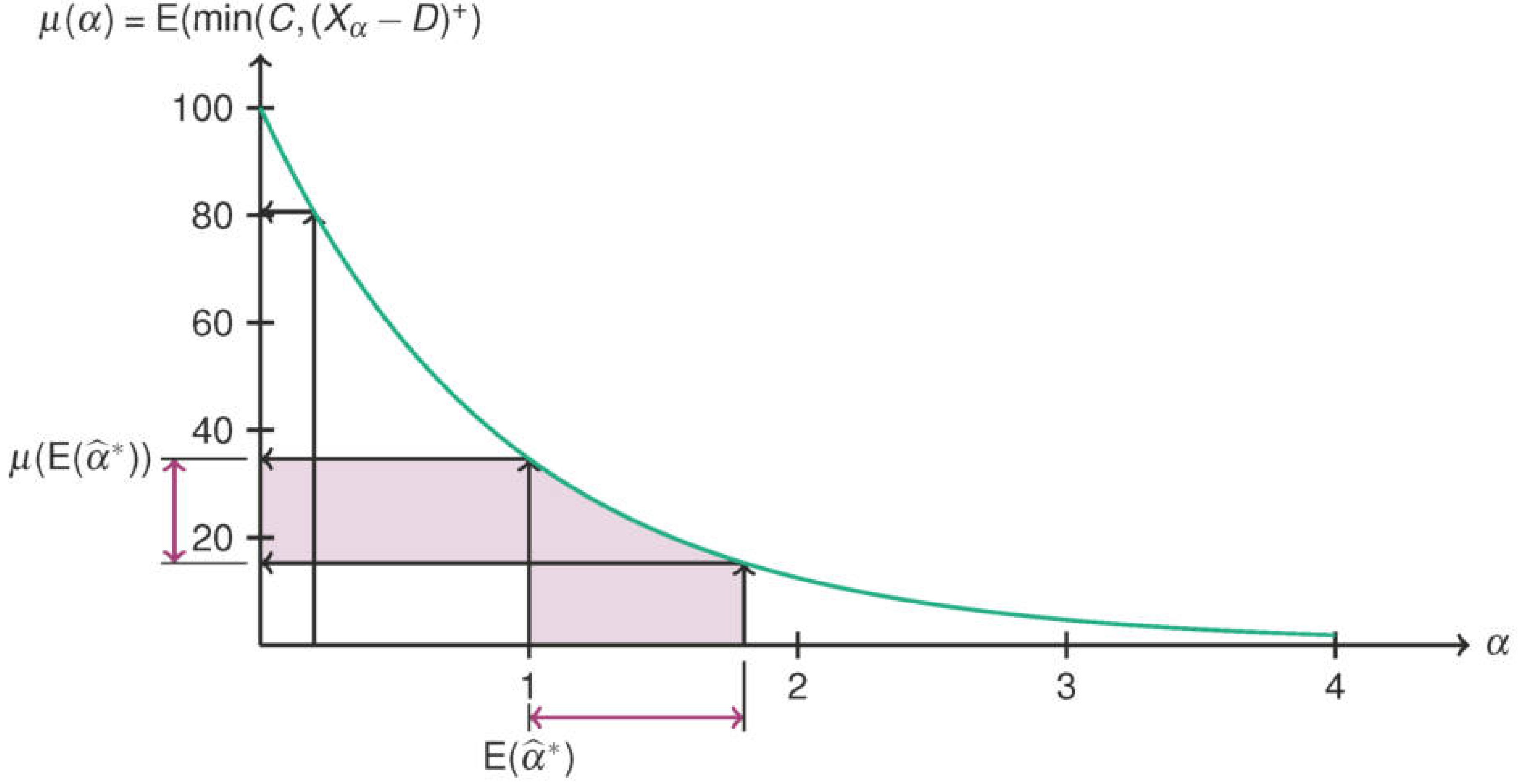
Jensen's Inequality



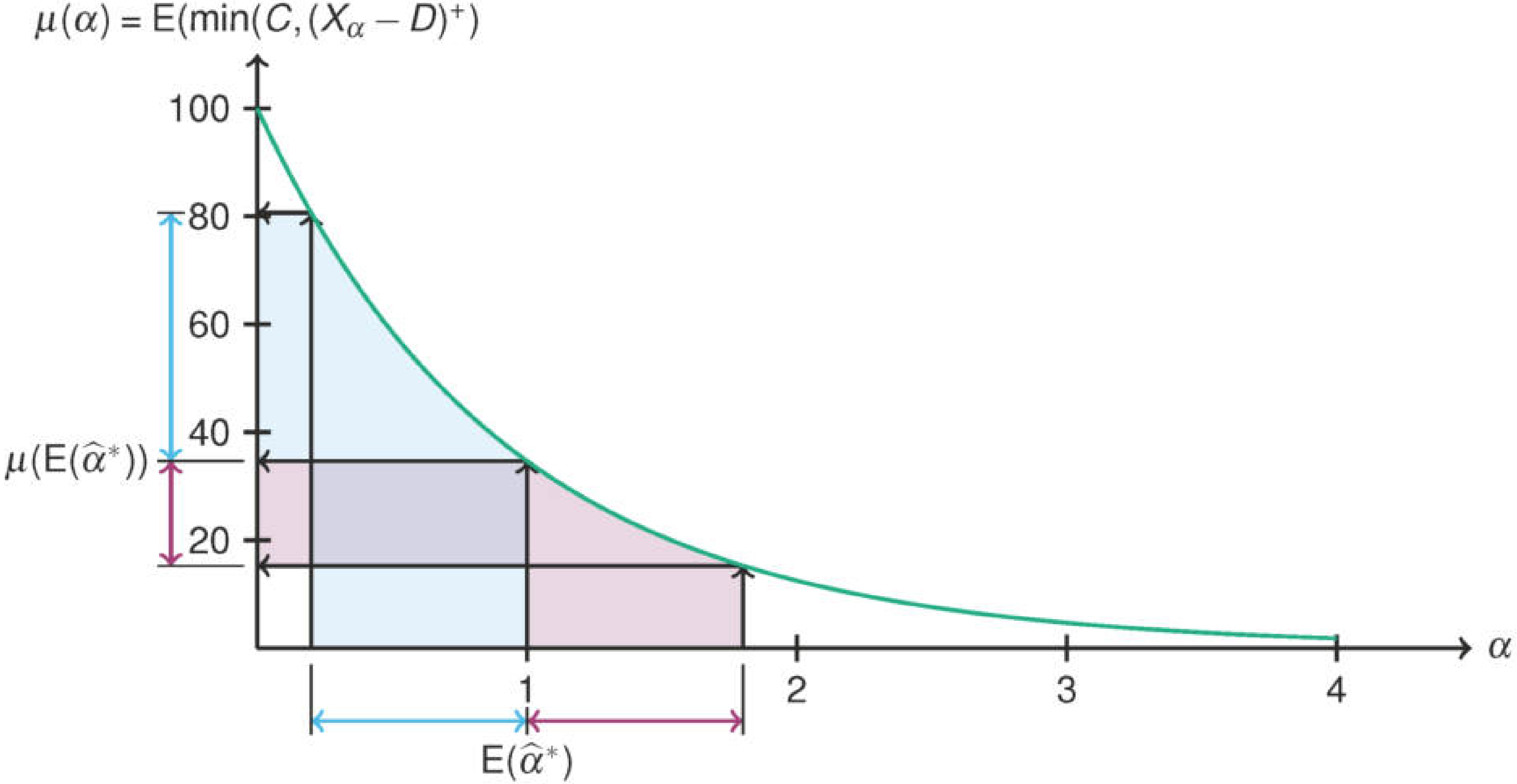
Jensen's Inequality



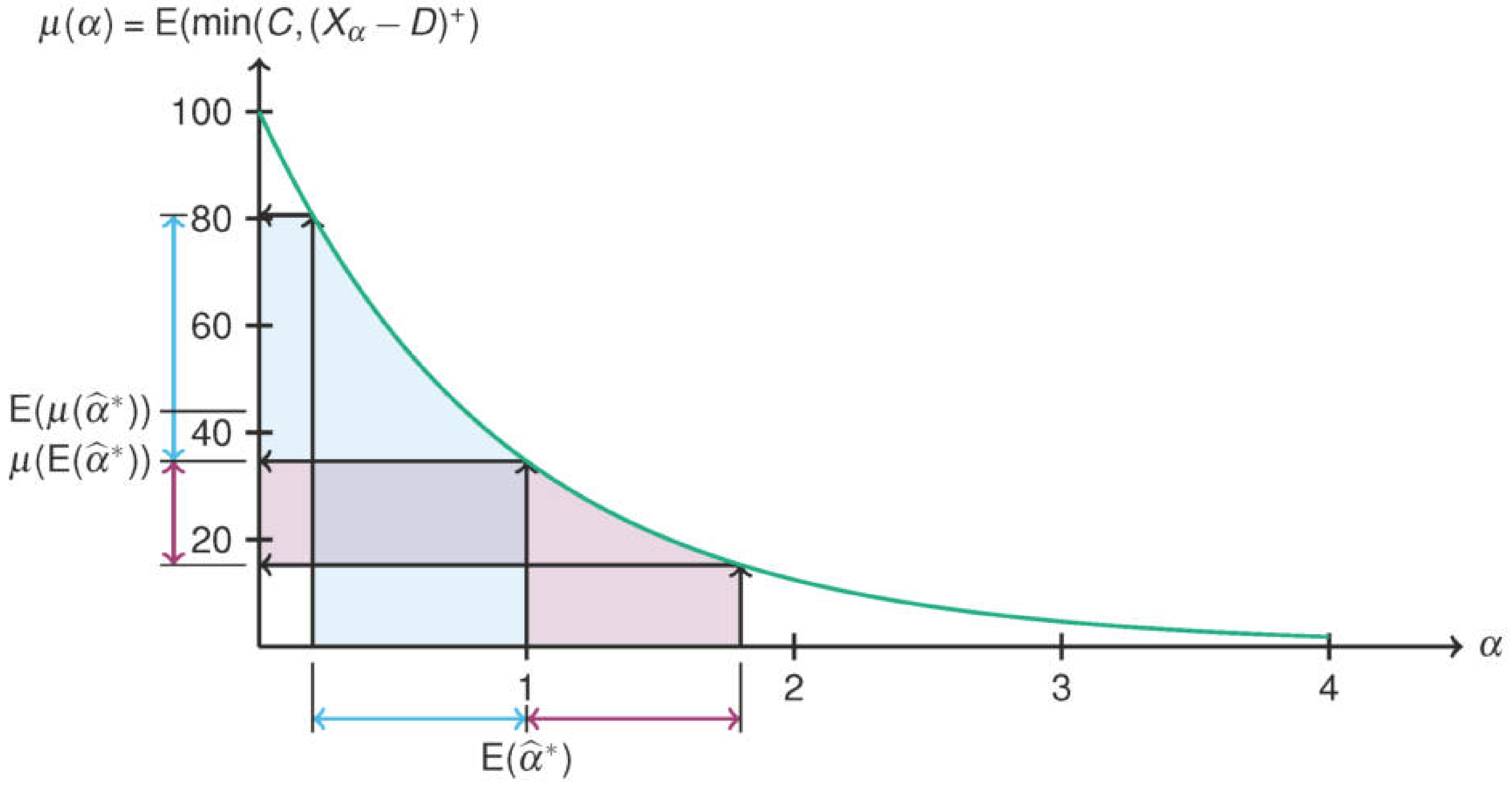
Jensen's Inequality



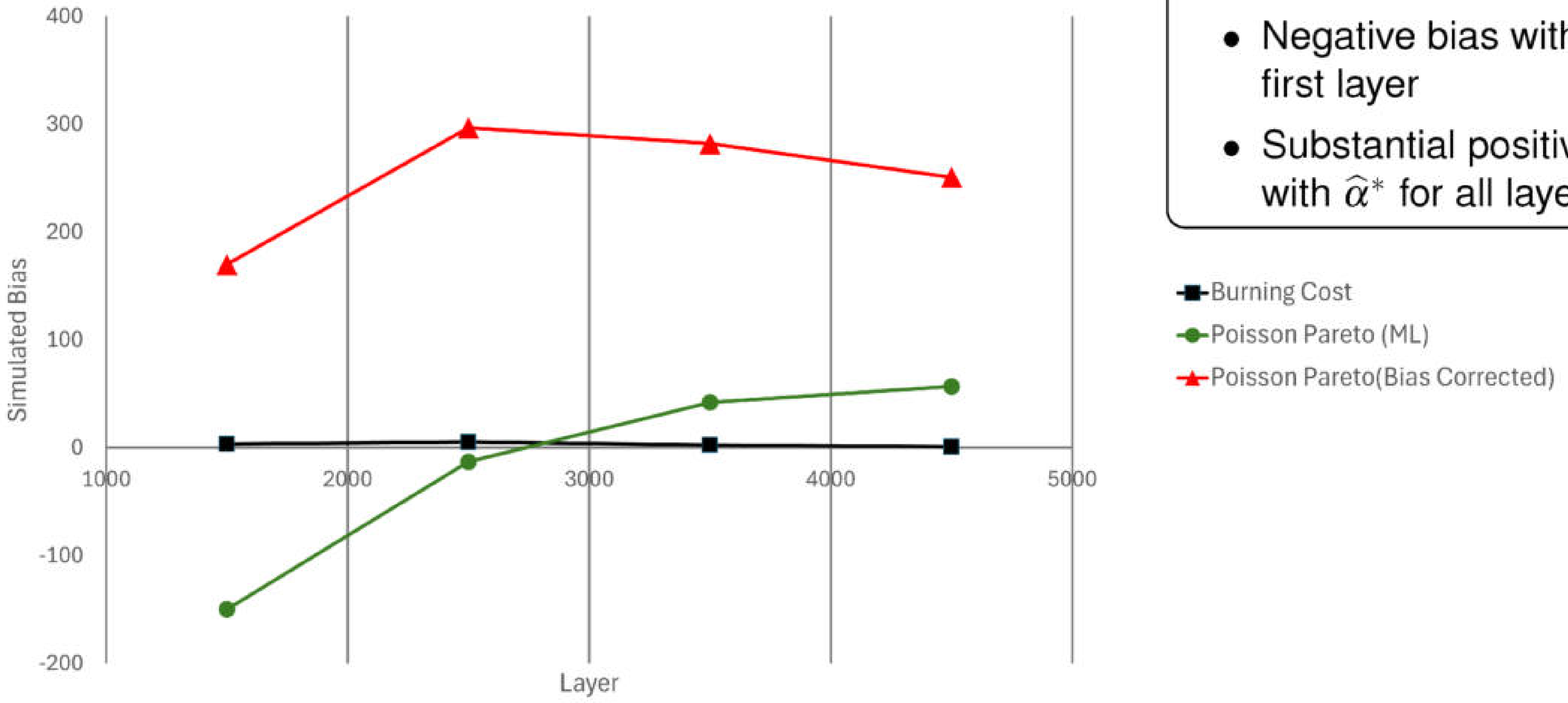
Jensen's Inequality



Jensen's Inequality

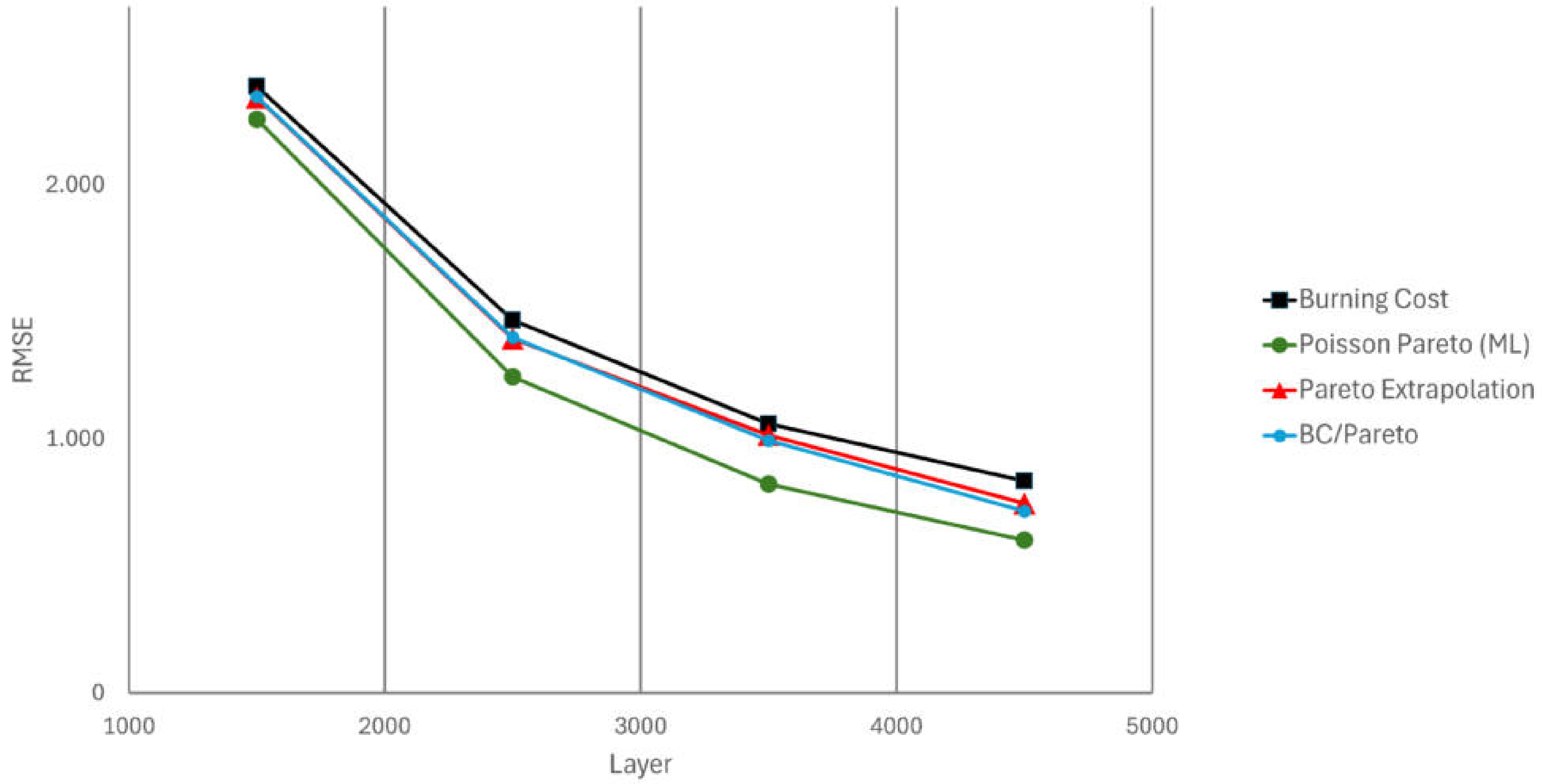


Bias in the case $\lambda = 5$, $\alpha = 2.0$

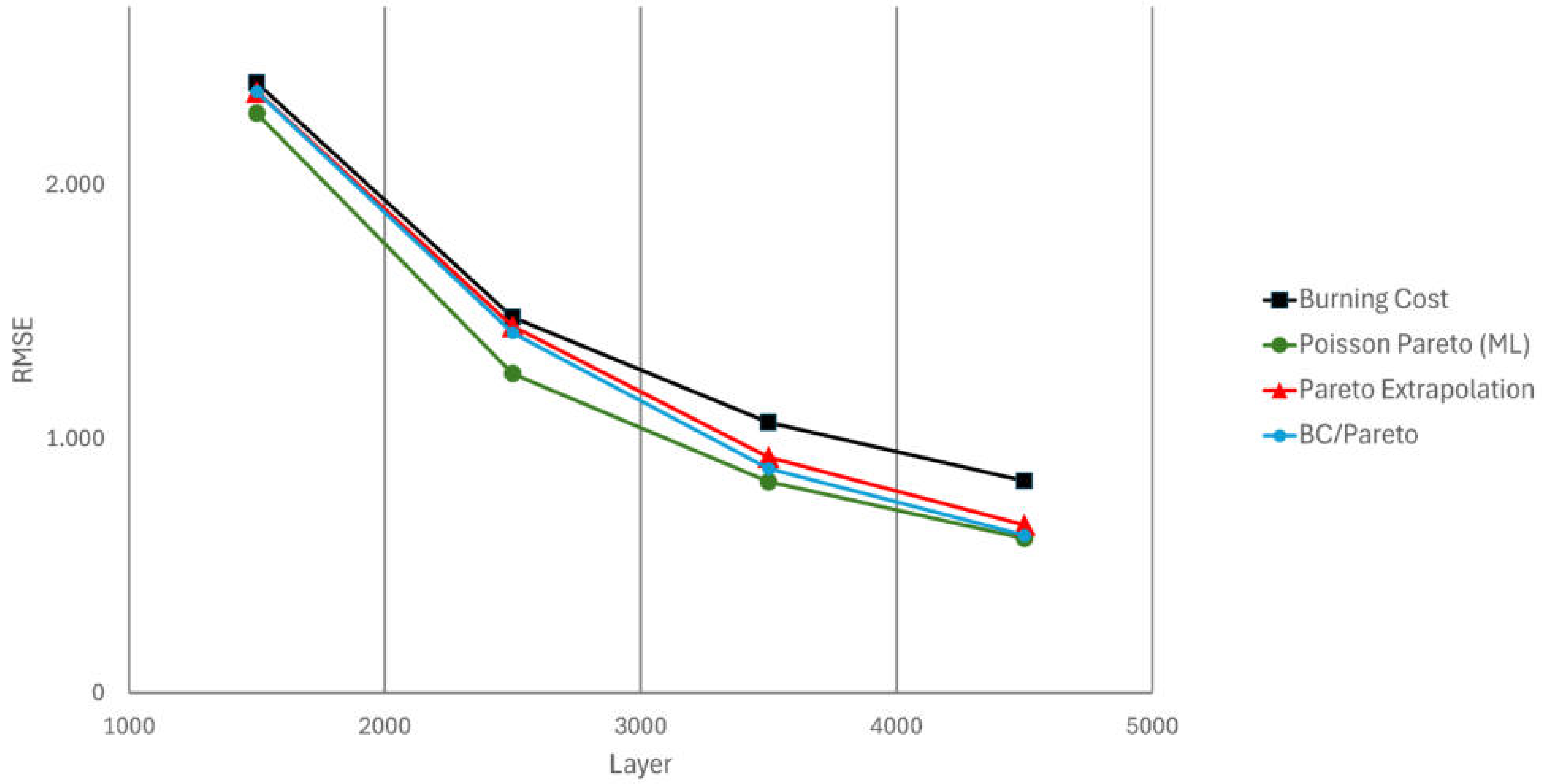


- BC has no bias
- Negative bias with $\widehat{\alpha}^{\mathrm{ML}}$ in first layer
- Substantial positive bias with $\widehat{\alpha}^{*}$ for all layers

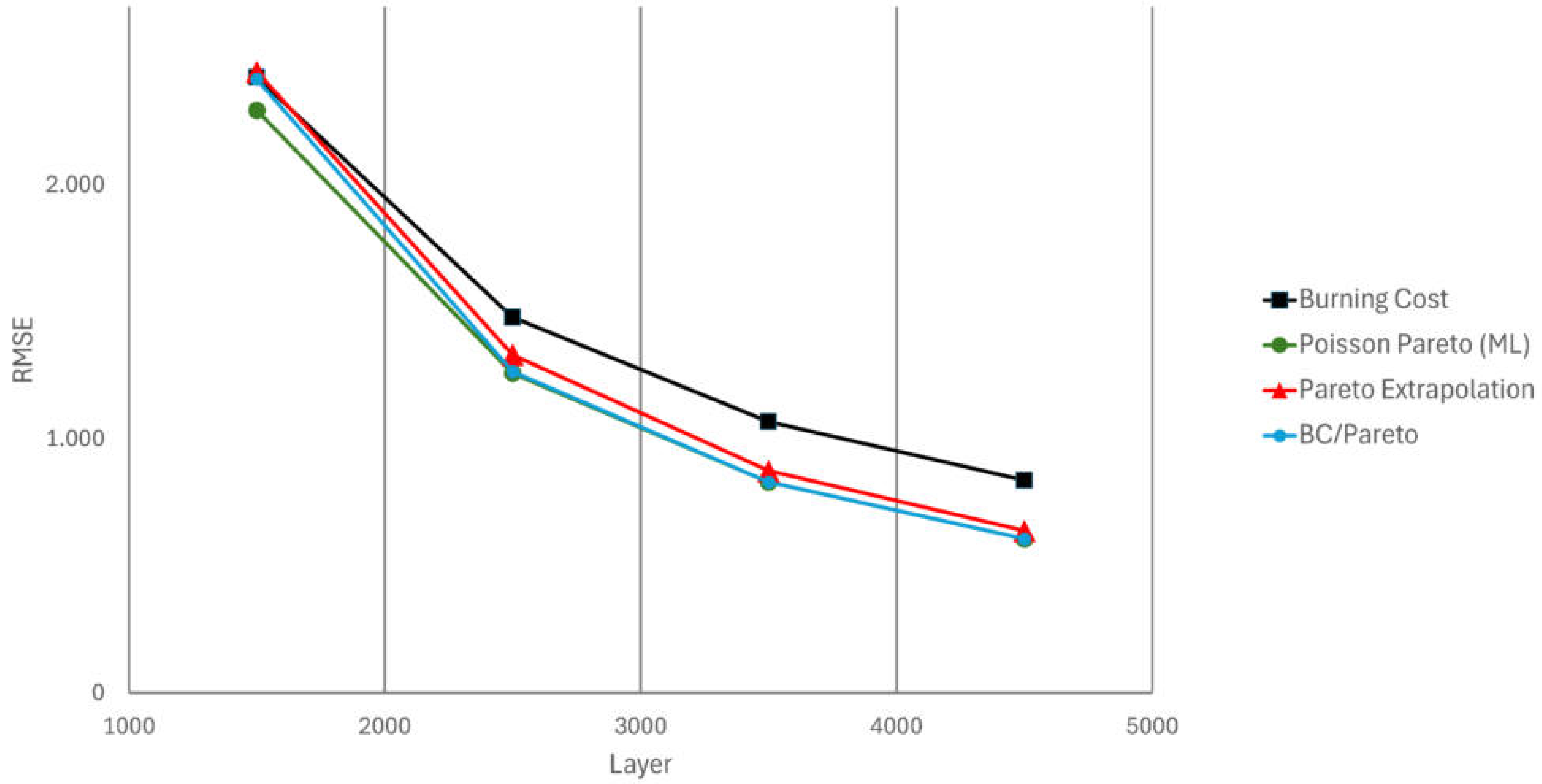
Burning Cost up to 3rd largest loss ($\lambda = 15$, $\alpha = 2.0$)



Burning Cost up to 5th largest loss ($\lambda = 15$, $\alpha = 2.0$)



Burning Cost up to 10th largest loss ($\lambda = 15$, $\alpha = 2.0$)



Observation 3

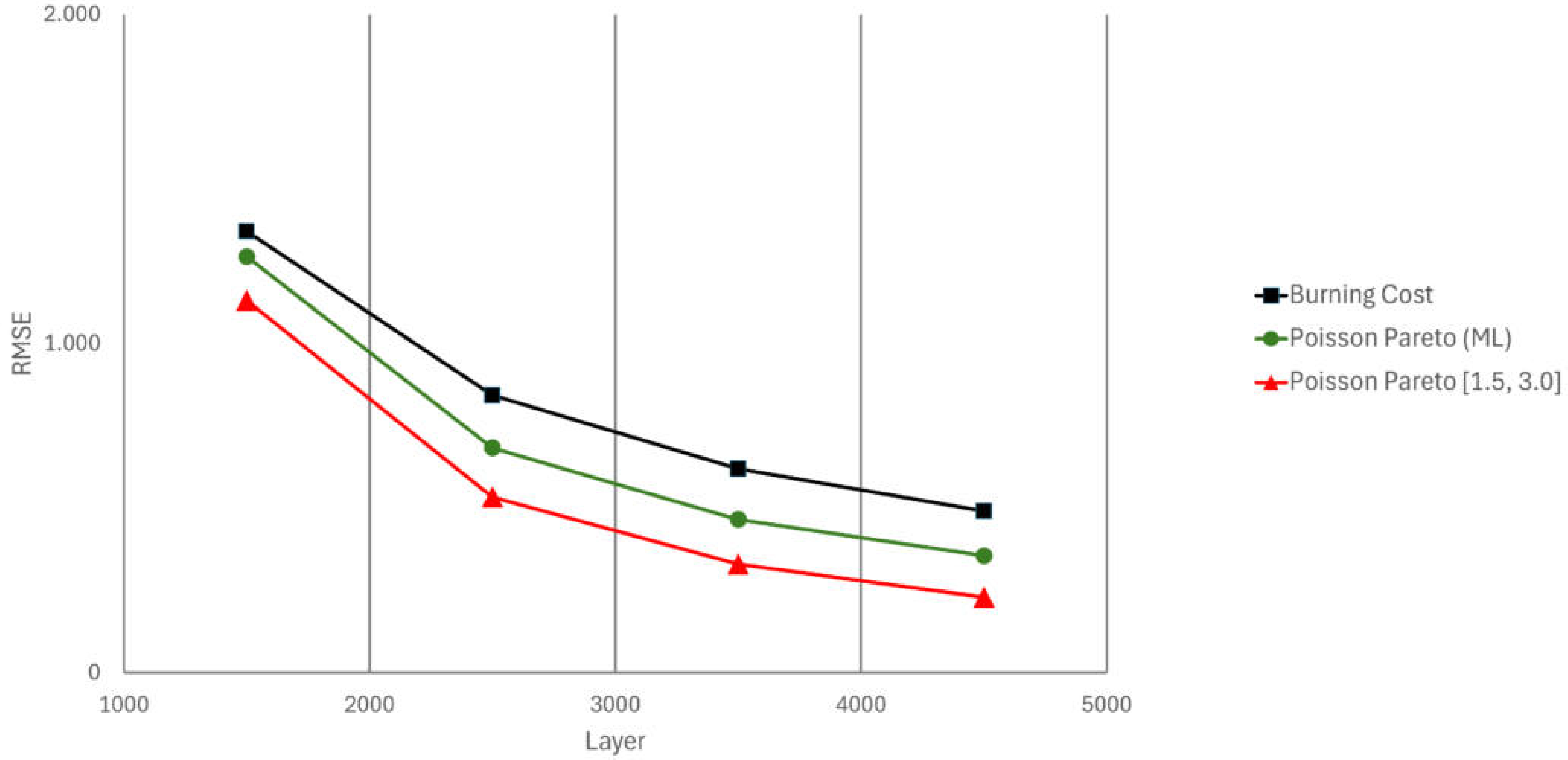
Methods that use the Burning Cost up to the k-th largest loss seem to have RSMEs between the Burning Cost and the Poisson/Pareto model.

Observation 3

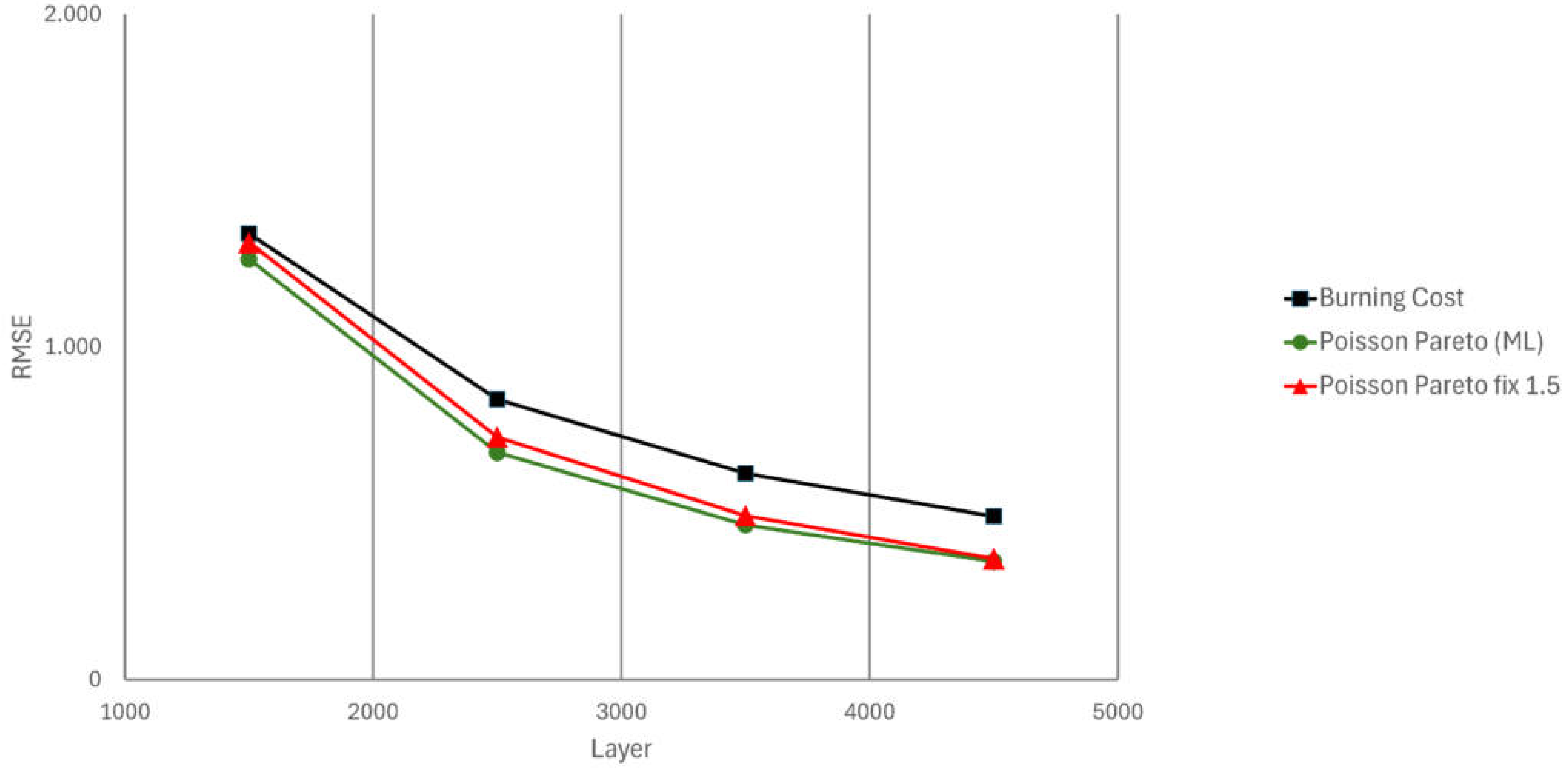
Methods that use the Burning Cost up to the k-th largest loss seem to have RSMEs between the Burning Cost and the Poisson/Pareto model.

Using the 3rd largest loss is closer to the Burning Cost model, using the 10th largest loss is close to the Poisson/Pareto model.

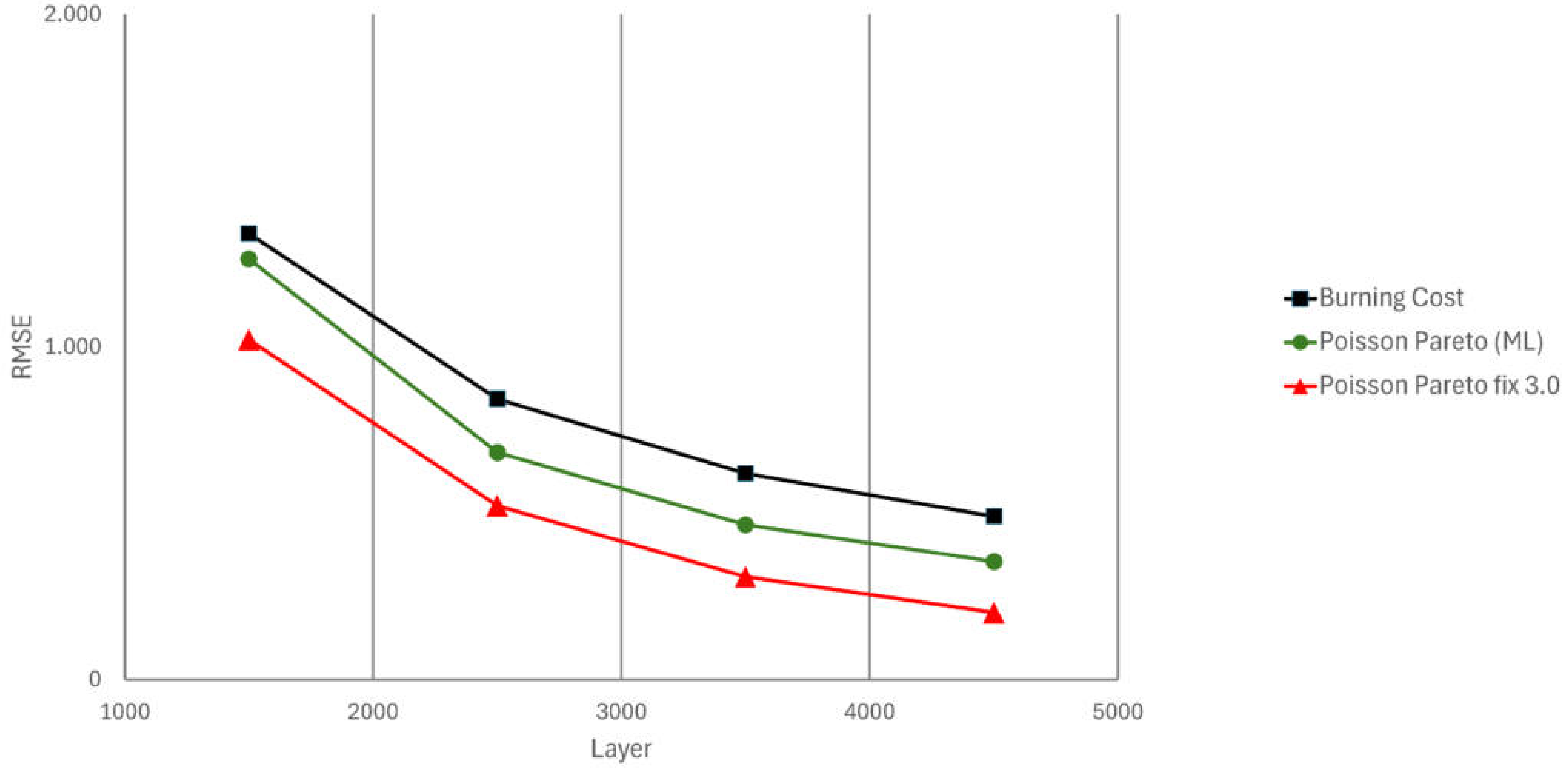
Use boundaries $\alpha_{\min} = 1.5$ and $\alpha_{\max} = 3.0$ for $\widehat{\alpha}^{ML}$ ($\lambda = 5$, $\alpha = 2.0$)



Use fixed *market alpha* $\alpha = 1.5$ ($\lambda = 5$, $\alpha = 2.0$)



Use fixed *market alpha* $\alpha = 3.0$ ($\lambda = 5$, $\alpha = 2.0$)



Observation 4

In case of a small λ the RMSE can be reduced substantially by using market knowledge for the estimation of the Pareto alpha:

- Upper/lower bounds for the estimator
- Market alpha
- Also a credibility approach (Bühlmann Straub) could be employed.

Observation 4

In case of a small λ the RMSE can be reduced substantially by using market knowledge for the estimation of the Pareto alpha:

- Upper/lower bounds for the estimator
- Market alpha
- Also a credibility approach (Bühlmann Straub) could be employed.

Note that the RSMEs are smaller, even if the portfolio's alpha deviates substantially from the market mean.

Observation 4

In case of a small λ the RMSE can be reduced substantially by using market knowledge for the estimation of the Pareto alpha:

- Upper/lower bounds for the estimator
- Market alpha
- Also a credibility approach (Bühlmann Straub) could be employed.

Note that the RSMEs are smaller, even if the portfolio's alpha deviates substantially from the market mean.

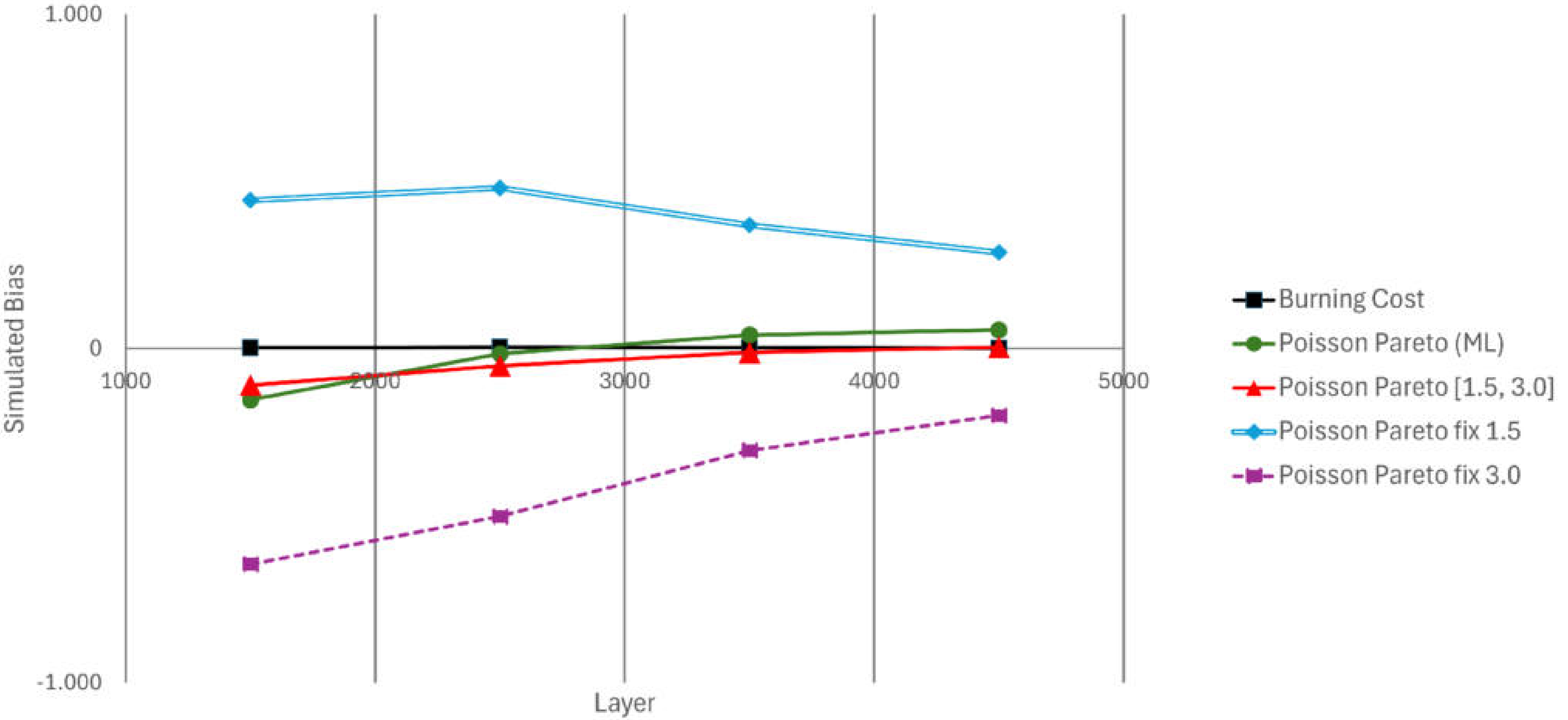
WARNING

Do not look exclusively at the RMSE!

Also care about the bias!

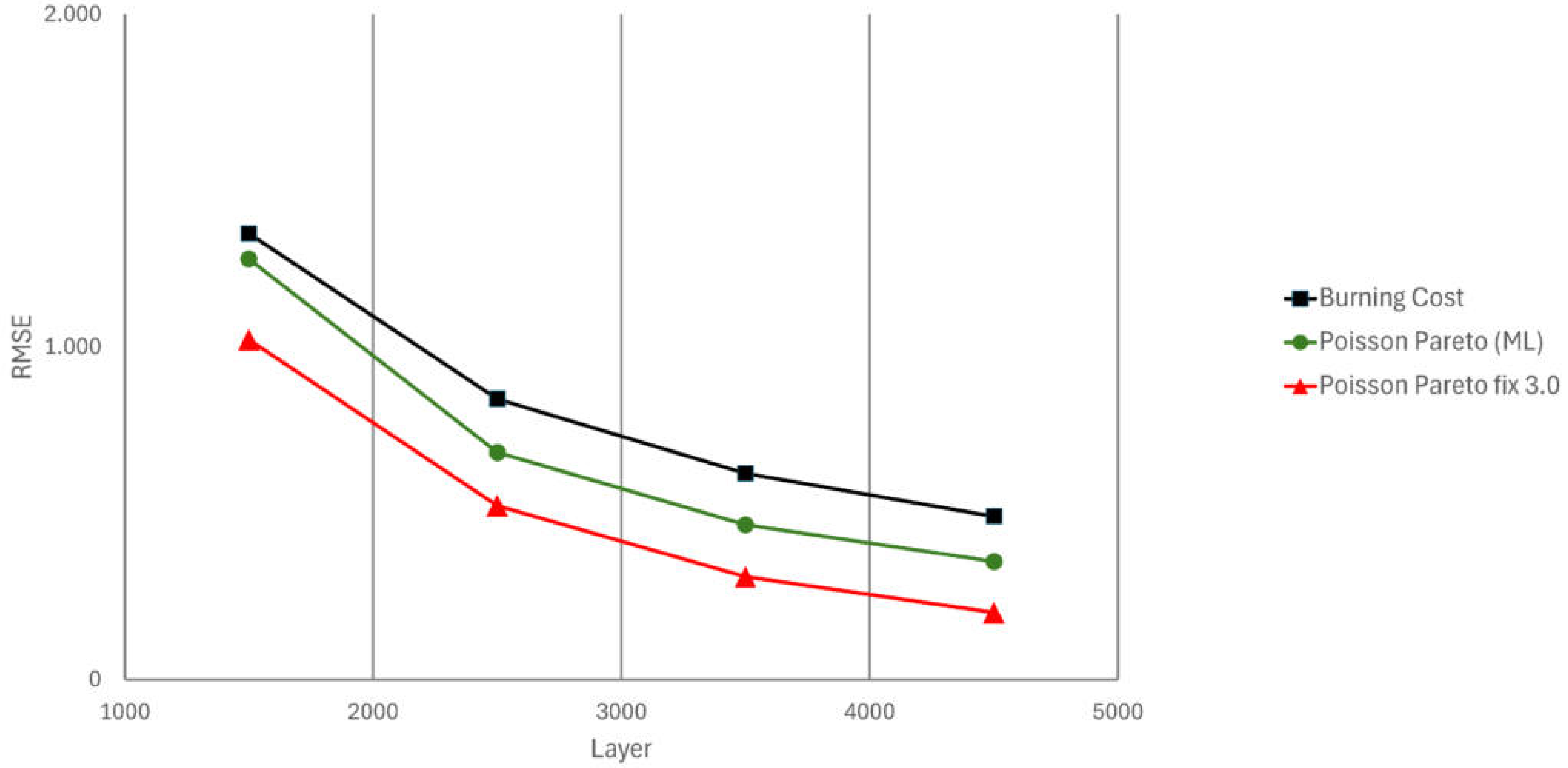
Bias with fixed/restricted alphas ($\lambda = 5$, $\alpha = 2.0$)

WARNING



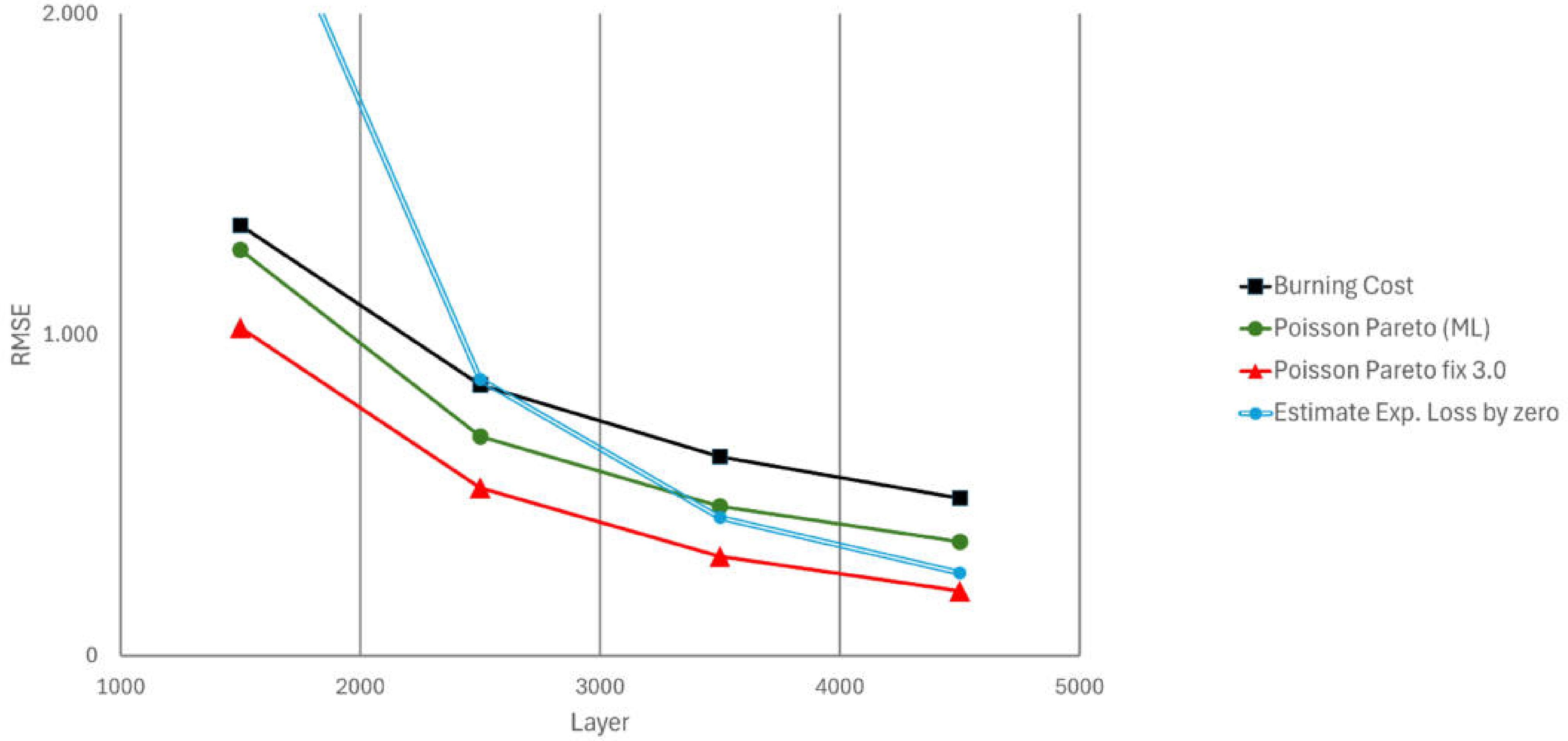
Use fixed *market alpha* $\alpha = 3.0$ ($\lambda = 5$, $\alpha = 2.0$)

WARNING



Use fixed *market alpha* $\alpha = 3.0$ ($\lambda = 5$, $\alpha = 2.0$)

WARNING



Agenda

1. Burning cost for per risk XLs
2. Pareto based pricing approaches
3. Simulation model for evaluating pricing approaches
4. Observations with underlying Pareto severity
5. Observations with other underlying severity distributions

Observation 5

Case 1: Underlying severity similar to Pareto

- If we just look at the RMSE, Poisson-Pareto still performed very good.

Observation 5

Case 1: Underlying severity similar to Pareto

- If we just look at the RMSE, Poisson-Pareto still performed very good.
- However, we observed a substantial bias in the first layer.

Observation 5

Case 1: Underlying severity similar to Pareto

- If we just look at the RMSE, Poisson-Pareto still performed very good.
- However, we observed a substantial bias in the first layer.
- Thus, methods using BC up to the k-th largest loss should be preferred.

Observation 5

Case 1: Underlying severity similar to Pareto

- If we just look at the RMSE, Poisson-Pareto still performed very good.
- However, we observed a substantial bias in the first layer.
- Thus, methods using BC up to the k-th largest loss should be preferred.

Case 2: Underlying severity deviates massively from a Pareto distribution

- Pareto based pricing methods can have a substantial bias.

Observation 5

Case 1: Underlying severity similar to Pareto

- If we just look at the RMSE, Poisson-Pareto still performed very good.
- However, we observed a substantial bias in the first layer.
- Thus, methods using BC up to the k-th largest loss should be preferred.

Case 2: Underlying severity deviates massively from a Pareto distribution

- Pareto based pricing methods can have a substantial bias.
- The observations with underlying Pareto severity do not apply in this case.

Vielen Dank für Ihre Aufmerksamkeit.

Fatima ezzahra Kherraz, Dr. Ulrich Riegel
fkherraz@glise.com, uriegel@munichre.com
+49 89 38912074, +49 151 51536526