# Title

Sustainability of wildlife harvest in stochastic social-ecological systems

# Author details

Elizabeth A. Law*[1], John D. C. Linnell[1], Bram van Moorter[1], Erlend B. Nilsen[1,2].

*Corresponding author: elizabeth.law@nina.no, workingconservation@gmail.com

[1] Norwegian Institute for Nature Research, PO Box 5685 Torgarden, NO-7485 Trondheim, Norway

[2] Faculty of Biosciences and Aquaculture, Nord University, Steinkjer, Norway

**Supporting information, code and data:**

- Supporting information: S1. MSE Model details. Provided as a .pdf file to be published alongside the article
- Supporting information: S2. Additional results. Provided as a .pdf file to be published alongside the article
- Code: Input data, simulation code and results are available in OSF repository DOI 10.17605/OSF.IO/U52RP (https://osf.io/u52rp/?view_only=e36abdca3e3c45d8813e6f7b20ce159a)
- Code: Analysis code and results are available in OSF repository DOI 10.17605/OSF.IO/CGWA6 (https://osf.io/cgwa6/?view_only=973dda4c88ea4a008c3b6e58ff149822)

Figures and tables, and their corresponding legends, are placed in text.

## Abstract

1. Sustainable wildlife harvest is challenged by complex and uncertain social-ecological systems, and diverse stakeholder perspectives. Heuristics could provide one avenue to integrate scientific principles and understand potential conflict in data-poor harvest systems. Management Strategy Evaluation (MSE) can be a useful tool to explore harvest options and implications from diverse perspectives, and aid in heuristic development.
2. We ran 176,910 stochastic simulation models to develop heuristics for sustainability in wildlife harvest systems. *Environmental contexts* included three simulated species distributed across the slow-fast life-history gradient (the great-unicorn, lesser-unicorn, and phoenix), two variability/uncertainty levels, and three starting population sizes. Optimal outcomes from four harvest strategies (constant, proportional, threshold-proportional, and threshold-increasing-proportional) were assessed under *evaluation contexts* reflecting multiple environmental, harvester, manager and societal sustainability objectives and ethical perspectives.
3. The results reveal fundamental challenges in obtaining sustainable outcomes in harvest systems: few scenarios produced good scores across all evaluation metrics and ethical perspectives. Composite evaluation metric sets and ethical perspectives strongly influenced perceived outcomes. Rawlsian ethical perspectives (considering the minimum score of multiple objectives) often revealed severe trade-offs between individual metrics, even when Utilitarian ethical perspectives (averaging scores of multiple objectives) view the same scenarios positively. Simple composite metrics popular in the theoretical literature often diverged from the holistic metrics that better reflect applied contexts.
4. Threshold and proportional systems performed better than constant harvest under Utilitarian ethics in 79-90% of cases, and 34-39% of cases with Rawlsian ethics. However, no strategy was optimal overall: each harvest system tested was near-optimal in at least one evaluation context in every environmental context.
5. *Synthesis and applications.* Given a lack of a singular optimum strategy, we recommend harvest systems should be chosen with clear reference to contextually appropriate metrics and ethics of interest when optimizing harvest systems for sustainability. Importantly, management

recommendations focused on maximizing harvest should be treated with scepticism if this is not explicitly identified as a key value for that socio-ecological system.

## Keywords

wildlife harvest, harvest protocol, population simulation, management strategy evaluation, socio-ecological systems, sustainable management, multiple objectives, uncertainty

## Introduction

Harvest is one of the most common forms of management for many wildlife species (DeVore, Butler, Wallace, & Liley, 2018; Riley et al., 2003), but achieving sustainability in wildlife harvest systems is challenging due to the complexity of social-ecological systems, with multiple uncertainties and diverse stakeholders (e.g. Mitchell et al., 2018; Gren, Häggmark-Svensson, Elofsson, & Engelmann, 2018). Wildlife harvest is important socially, culturally and economically for both direct benefits (e.g. meat, income, recreation, tradition) and avoiding costs and human-wildlife conflicts (e.g. vehicle collisions, predation on domestic animals, and competition or pathogen spread between wild and domestic stock; DeVore et al., 2018; Gren et al., 2018; Linnell et al., 2020, 2020; Mitchell et al., 2018). Wildlife-harvest systems are typically managed with an overarching aim of sustainability (Weinbaum, Brashares, Golden, & Getz, 2013), yet 'sustainability' is a multi-faceted but ill-defined term (Quinn & Collie, 2005) often poorly applied in practice (Weinbaum et al., 2013). Definitions, while centring on ensuring persistence of the species and its harvest, contemporarily encompass diverse economic and social concepts, ecological, habitat, and ecosystem-based criteria, and precaution under uncertainty (Hilborn et al., 2015; Quinn & Collie, 2005).

Despite established theory on optimal harvest strategy (e.g. Hilker & Liz, 2020; Lande, Engen, & Saether, 1994, 1995; Lande, Sæther, & Engen, 1997; Sæther, Engen, & Lande, 1996), in practice determining quotas in terrestrial systems is often an inexact, adaptive science at best (Artelle et al., 2018). Due to limited resources and poorly developed institutional frameworks, many wildlife management systems lack all but the most rudimental parameters (van Vliet & Nasi, 2019; Weinbaum et al., 2013), and even in the best cases elements of social-ecological systems remain uncertain or contested (Bischof et al., 2012; Corlatti, Sanz-Aguilar, Tavecchia, Gugiatti, & Pedrotti, 2019; Nilsen, 2017; Pellikka, Kuikka, Lindén, & Varis, 2005; Stevens, Bence, Porter, & Parent, 2017). From fisheries management systems, literature syntheses suggest strong context-dependencies of optimal strategies (Deroba & Bence, 2008), but no such synthesis has been conducted for terrestrial systems. In many cases, terrestrial wildlife harvest management simply lacks science and transparency (Artelle et al., 2018; Weinbaum et al., 2013). This opens the door for political intervention in quota setting, exposing management to potential social and legal conflict (Artelle et al., 2018).

Sustainability may be improved through the use of heuristics and simulation models. Heuristics are practical and accessible guidelines designed to give good 'rules-of-thumb' (i.e. good outcomes over a wide range of cases) in applied management scenarios where more detailed information is lacking (Leung, Finnoff, Shogren, & Lodge, 2005). Heuristics can be derived from empirical experience, or deduced from simulation models (Davis, Chadès, Rhodes, & Bode, 2019; Deroba & Bence, 2008). Simulation models help to formalise knowledge, and are well established in conservation and wildlife-management contexts. Typically these focus on stochastic population dynamics, for example in population viability analysis (Lacy, 1993; Miller, Furness, Trinder, & Matthiopoulos, 2019; Weinbaum et al., 2013), while traditional harvest models couple this with harvest (Hilker & Liz, 2019; Lande et al., 1995; Sæther et al., 1996). Management Strategy Evaluation (MSE) models expand from these, encompassing stochastic simulations of management in socio-ecological systems incorporating a more holistic set of ecological and social components (Bunnefeld, Hoshino, & Milner-Gulland, 2011). MSE models are well established in fisheries (Punt, Butterworth, Moor, Oliveira, & Haddon, 2016) and increasingly used in terrestrial management scenarios (e.g. Bled & Belant, 2019; Eriksen, Moa, & Nilsen, 2018; Manning, Stevens, & Williams, 2019; Miller et al., 2019; Mitchell et al., 2018; Riley et al., 2003). MSE models address key knowledge gaps regarding the implications of uncertainty in the multiple socio-economic facets of wildlife harvest systems (Gren et al., 2018), and allow levels of systematic assessment impossible in real-world experiments. Heuristics developed from MSE models may be able to address the science-policy gap between theoretical harvest models and real-world application of harvest strategies, through 1) improving our understanding of more complex and uncertain socio-ecological systems, and 2) shifting the focus from what strategies are optimal in constrained theoretical settings to what is likely to be acceptable in a diverse range of environmental and social evaluation settings, including under different ethical perspectives.

Consideration of diverse ethical perspectives as to what is valued, and how different values are appreciated, is a growing focus of environmental management (Friedman et al., 2018). This facilitates social equity by better representing diverse stakeholder values and perspectives – a virtuous social outcome in itself as well as contributing to the success and sustainability of management actions (Friedman et al., 2018; Law et al., 2018). Utilitarian ethics emphasise aggregate utility ('the greatest good for the greatest number'), and are commonly applied via summation or averaging over a set of outcomes (for example, via cost-benefit evaluations; Law et al., 2018). However this ethic is not universally held and is criticised for allowing concerns of the majority to overwhelm concerns of minorities (Wilson & Law, 2016). In contrast, Rawlsian ethics focus on improving the outcomes for the stakeholders that fare the worst, typically represented through maximising the minimum score of a set of outcomes (i.e. a maxi-min function; Rawls, 1971). People display both Utilitarian and Rawlsian ethics when making personal decisions (Kameda et al., 2016). However, despite Rawlsian patterns being more common when these decisions affect others (Kappes, Kahane, & Crockett, 2016), the environmental decision-making literature has tended to be dominated by Utilitarian 'cost-benefit' or aggregate sum metrics (Friedman et al., 2018; Law et al., 2018). Further ethical perspectives (not well captured by either Utilitarian or Rawlsian functions) are gaining popularity in environmental management, for example concerns for animal welfare, animal rights, and 'compassionate' conservation (Hampton, Warburton, & Sandøe, 2019; Hayward et al., 2019). An understanding of alternative stakeholder perspectives is of practical importance for understanding the level of satisfaction that alternative stakeholders may have under different decisions, and consequently how contested or sustainable decisions may be.

To develop heuristics for sustainable wildlife harvest, we construct MSE models within a consistent model framework, spanning diverse environmental contexts including a gradient of species life-history types, uncertainty, and starting conditions. We simulate a set of species from across the fast-slow life-history gradient, a commonly used motif for theory development in wildlife phenomena describing patterns of covariation in life-history traits across body size, longevity, and fecundity (Bielby et al., 2007; Williams, 2013). We evaluate sustainability over a range of metrics and ethical perspectives relevant for terrestrial contexts, and include multiple types of variability representing both temporal stochasticity and parameter uncertainty (McGowan, Runge, & Larson, 2011), in resource, monitoring, management decision, and harvest implementation components. We compare the simulations to uncover: 1) which strategies are optimal in different contexts? 2) which strategies give acceptable outcomes across a diverse range of contexts? 3) what simple heuristics regarding environmental and evaluation contexts can be developed?

## Materials and methods

We develop a MSE model that generalises a terrestrial wildlife-harvest system, with components of 1) resource dynamics, 2) monitoring observations, 3) quota setting, 4) harvest implementation, and 5) sustainability evaluation. Simulations occur in yearly time steps ($t$), across a time series of 20 years (broadly considered long term for applied management plans), with multiple iterations ($i$ = 1000) per scenario. Full model description and parameter values are available in Supporting Information S1, and only summarised here.

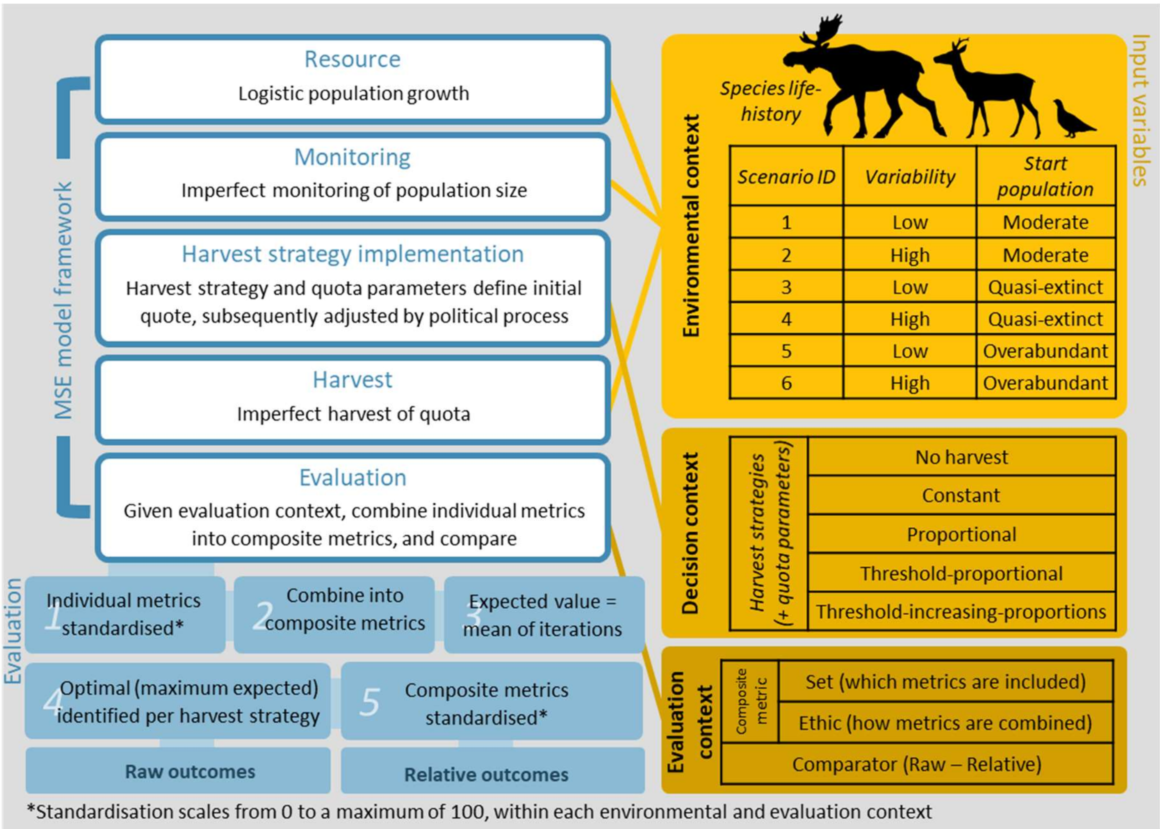
### MSE framework

The **resource component** simulates growth of a population $N_{i,t}$, using logistic growth determined by the population's intrinsic growth rate, $r_{i,t}$, and carrying capacity, $K$. The **monitoring component** is simulated by a single variation factor ($m_{i,t}$) acting on $N_{i,t}$, to give an estimate of the population size ($\widehat{N_{i,t}}$), to be used as the basis for management decisions. The **management-decisions component** comprises two parts. First, a harvest strategy is applied, converting $\widehat{N_{i,t}}$ into an initial quota, $Q_{i,t}$, given a set of quota parameters. $Q_{i,t}$ is then subject to random variation ($q_{i,t}$) to simulate the political interference in the quota setting process, to give a modified quota $Q'_{i,t}$. The **harvest implementation component** simulates imperfect harvest implementation, effected as a proportional variation ($h_{i,t}$) around $Q'_{i,t}$ to give the harvest ($H_{i,t}$). This amount is then removed from $N_{i,t}$, before continuing to the next timestep. Stochastic parameters include $r$, $m$, $q$, and $h$, which simulate environmental stochasticity, imperfect implementation, and parameter uncertainty, using normal distributions partitioned over years ($t$) and iterations ($i$).

The **evaluation component** occurs after each simulation is complete, calculating performance metrics of each iteration over the entire timeframe, and summarising over iterations in the scenario run (see details below and in Supporting Information S1). Individual evaluation metrics reflect different stakeholder concerns over various socio-ecological and harvest-based sustainability objectives, and are summarised into composite metrics under alternative evaluation contexts (i.e. with different emphases and ethics; Table 1,2).

## [Figure 1: MSE framework]

**Figure 1:** The Management Strategy Evaluation (MSE) model simulates a wildlife harvest system over a 20 year timeframe, with each environmental and decision scenario including 1000 stochastic iterations. Evaluation combines individual metrics into composite scores, using different functions to simulate different evaluation contexts. Species types span a fast-slow life-history gradient, determining growth rates and carrying capacity, variation levels in growth rates and monitoring variability, and critical thresholds. Stochastic parameters simulate yearly stochasticity and iteration level uncertainty. A full description of the model and parameter values are specified in Supporting Information S1.




## Environmental context and decision variable parameters

Species life-history, variability/uncertainty, and starting population scenarios collectively represent the **environmental context**. We simulate three virtual species spanning a slow-fast life-history gradient of common game species (Table S1.1), based on wildlife harvested in a Norwegian context but with global relevance. The great-unicorn resembles a large ungulate (e.g. moose, *Alces alces*), and is assumed to have a low growth rate, carrying capacity, monitoring variation, and critical thresholds for evaluating population size. The lesser-unicorn resembles a small ungulate (e.g. roe deer, *Capreolus capreolus*), with a moderate growth rate, carrying capacity, monitoring variation, and critical thresholds. The phoenix is reflective of a game bird (e.g. willow ptarmigan, *Lagopus lagopus*), with a large potential growth rate, carrying capacity, monitoring variation, and critical thresholds.

For each species we simulated two variability scenarios, where variability in $r$, $m$, $q$, and $h$ was *low* or *high*, and three starting populations: 1) the midpoint of low and high critical thresholds (*moderate*), 2) *quasi-extinction*, and 3) *overabundance*. Alternative starting populations test the robustness of the harvest strategies to extreme perturbations in population size, as well as being relevant for special management cases (e.g. overabundant species, or recovery of endangered species into harvestable populations). Variability and starting population scenario combinations are identified numerically (1-6) defined in Figure 1.

The harvest strategies and quota parameters represent **decision variables**. Harvest strategies analysed include '*constant*' (a set number of individuals harvested yearly), '*proportional*' (a set proportion of the population harvested yearly), '*threshold-proportional'* (a set proportion taken yearly, provided the population is above a certain threshold), and '*threshold increasing-proportions*' (provided the population is above a certain threshold, the proportion taken increases as the population size increases). We assume that the harvest strategies and associated parameters (constants, thresholds, and proportions) remain consistent throughout the timeframe (note that the resulting quota adapts to the population size in all but the constant harvest). We employed a grid search method across a wide range of possible quota parameter options in order to identify and compare optimal strategies across a diversity of potential objectives (see Table S1.2).

## Evaluation contexts and comparisons

**Evaluation contexts** are designed to reflect different stakeholder perspectives on outcomes from the simulations. These determine which individual metrics are of interest (i.e. *the composite metric set*; Table 1,2), how they are summarised (i.e. *the composite metric function*, reflecting alternative ethical perspectives), and to which other outcomes a comparison is being made (i.e. the comparator).

We assess six *composite sets* with varying emphasis and degrees of complexity (Table 2). These range from a *complete* set, including all metrics, to a *classic* set that are comprised of only those metrics commonly seen in the classic theoretical literature (namely maximize harvest and persistence). Others represent particular contexts, such as focus only on *population* or *harvest* related metrics, or all except overabundance as this may be of low concern in some contexts (e.g., for *complete small-game*). We combine the elements of each composite set using two *composite functions*: maximizing a weighted mean score representing a Utilitarian (aggregate benefit) ethic, and maximizing the minimum score from the set representing a Rawlsian (maxi-min) ethic. Individual metric scores are first standardised (scaled so that 100 represents the most desirable expected outcome possible, e.g. the largest probability of non-extinction, or the largest mean harvest) over decision variables for each respective environmental context before combination. Because individual metric scores within iterations are not independent, composite metric scores were calculated for each iteration, before being summarised over the decision variables (Supporting Information S1). Composite metric scores therefore represent outcomes as perceived under specific ethical and stakeholder contexts.

Comparative analysis focused on the optimal outcomes for each harvest strategy: each harvest strategy was represented by the score from the quota parameters that maximized the expected outcome (i.e. mean across the 1000 iterations) under each environmental and evaluation context. To reflect how satisfied stakeholders may be with optimized outcomes with respect to that harvest system only, we assessed **raw scores** for each composite metric (where 100 represents perfect scores across all individual metrics in the composite set). To show relative optimality of the harvest strategy in that environmental and evaluation context, we assessed **relative scores** (where 100 represents the best score achieved across decision variables, i.e. all harvest systems and quota parameters, within each respective environmental context). In scaling the relative scores, all harvest systems achieving the best score gained a score of 100, even if this 'best' score was zero. Rawlsian scores can also indicate the minimum potential for trade-offs to occur, in that they give the maximal minimum score from the set. However, trade-offs could be even worse than indicated by Rawlsian scores if optimization for this metric is not achieved, for example if a utilitarian ethic is used, actors are self-serving, power is unequal, or if actors are malevolent and actively seek to minimise the outcomes of others.

## Heuristics

We defined 'heuristics' as a set of simple rules or guidelines for a) choosing an optimal harvest system, and b) when contextual factors are likely to give 'good' (but not necessarily optimal; arbitrarily defined as scoring 85-100), or 'better' (in the case of pairwise comparisons) perceived outcomes. We sought heuristics via plotting outcome scores and ranking strategies for different contexts, and developing decision trees for optimal strategies and the likelihood of good outcomes being perceived. The factorial design of the simulations also allowed us to assess the pairwise differences by matching outcomes from the different contextual factors, all other variables held constant. We excluded the 'no harvest' strategy from these comparisons.

We constructed decision trees based on conditional inference methods (Hothorn & Zeileis, 2015): binary recursive partitioning using regression relationships, first testing if there are any significant relationships of the predictors to the response variable, and then, if so, implementing the binary split with the strongest association with the response variable, and repeating until no further significant relationships are found. These have the benefit of being easily interpretable, limiting recursive partitioning at reasonable levels, and have reduced bias for mixed variables (Strasser & Weber, 1999).

We constructed the model in R (R Core Team, 2020), using tidyverse (Wickham, Averick, et al., 2019) and truncnorm (Mersmann, Trautmann, Steuer, & Bornkamp, 2018), parallelized with doSNOW (Microsoft Corporation & Weston, 2019). For the decision trees, we used default methods under partykit::ctree (Hothorn, Hornik, & Zeileis, 2006; Hothorn & Zeileis, 2015). For graphics, we used ggplot2 (Wickham et al., 2020), ggtable (Wickham, Pedersen, & RStudio, 2019), ggparty (Borkovec et al., 2019), and cowplot (Wilke, 2019).

[Table 1: Individual sustainability metrics]

**Table 1:** Sustainability metrics represent a wide variety of common stakeholder concerns, and include fundamental sustainability objective of non-extinction, as well as other *population-based* and *harvest-based* metrics. Here they are constructed so that within each metric higher scores are more desirable.

| Objective group | Objective | Criteria | Code |
|---|---|---|---|
| *Persistence* | **Avoiding extinctions.**<br>A fundamental objective of ecological and economic sustainability. | 1 – Probability population goes extinct by year 20 | *probability of non-extinction* |
| *Population* | **Population stability.**<br>Avoiding population extremes. | Number of years population remains between *high* and *low* critical thresholds | *stable population* |
| | **Avoiding low or functionally extinct populations.**<br>To provide adequate populations for harvest, ecological functionality, and buffer against extinctions. | Number of years population remains above the *quasi-extinction* critical threshold | *above quasi-extinct* |
| | | Number of years population remains above the *low* critical threshold | *above low* |
| | **Avoiding high and overabundant populations**.<br>To minimize wildlife conflict and ecological damage from overabundant populations. Note, this may not be a concern for small game species. | Number of years population remains below *high* critical threshold | *below high* |
| | | Number of years population remains below the *overabundance* critical threshold | *below overabundant* |

| | | | |
|---|---|---|---|
| *Harvest* | **Mean annual harvest.**<br>To provide the maximum opportunity for economic and social benefits of harvest. | Mean yearly harvest | *harvest mean* |
| | **Minimum harvest** experienced across the timeframe.<br>To maximize harvest opportunity over every point in the timeframe. | Minimum harvest size across the timeframe | *harvest minimum* |
| | **Avoiding years experiencing zero harvest.**<br>To provide consistency of harvest experience and income for harvesters and associated economies. | Number of years harvest is not zero | *harvest non-zeros* |
| | **Limiting harvest variability.**<br>While some variability may be accepted as an inevitability in variable contexts, consistency of harvest improves predictability and the consistency of capital required for its implementation. | 0 – Standard deviation of harvests over the timeframe | *harvest consistency* |

## [Table 2: Composite metrics]

**Table 2:** Composite metrics are comprised of six different *sets* of individual metrics, combined using two different *functions* to reflect alternative ethical perspectives. Inclusion in sets is denoted by a tick (included) or cross (not included), and the assigned weights for Utilitarian function shown in brackets.

| | **Individual metric** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Persistence* | *Population* | | | | | *Harvest* | | | |
| **Composite metric set** | Probability of non-extinction | Above quasi-extinct | Above low | Stable population | Below high | Below overabundant | Harvest mean | Harvest minimum | Harvest non-zeros | Harvest consistency |
| Complete | ✓ (*1*) | ✓ (*0.2 each*) | | | | | ✓ (*0.25 each*) | | | |
| Population focus | ✓ (*1*) | ✓ (*0.2 each*) | | | | | ✗ | | | |
| Harvest focus | ✓ (*1*) | ✗ | | | | | ✓ (*0.25 each*) | | | |
| Complete (small game) | ✓ (*1*) | ✓ (*0.5 each*) | | ✗ *(no concern for overabundance)* | | | ✓ (*0.25 each*) | | | |
| Classic pop.+harv. | ✓ (*1*) | ✗ | | ✓ (*1*) | ✗ | | ✓ (*1*) | ✗ | | |
| Classic harv. | ✓ (*1*) | ✗ | | | | | ✓ (*1*) | ✗ | | |

| **Composite metric function** | | |
|---|---|---|
| *Ethic* | Utilitarian (maximize aggregate good) | Weighted mean of included metric scores (assuming equivalent emphasis on persistence, population, and harvest groups) |
| | Rawlsian (maximize minimum outcome) | Minimum score of included metric scores (all individual metrics weighted equally) |

## Results

Composite metric scores varied over quota parameter options within each harvest strategy. Suboptimal harvest strategies with optimized quota parameters performed better than optimal strategies with poorly selected quota parameters (Figure 2). Constant harvest strategies, 'faster' life-histories, and more variable environmental contexts had greater outcome uncertainties (Supporting Information S2.1). Steep declines in performance occurred with overharvesting under constant and proportional strategies without thresholds. While such risks are likely to be a consideration in applied decision contexts, the following results are based on expected (mean) outcomes from each set of iterations, and are therefore representative of risk-neutral decision-making only.

Our simulations show that there was no single optimum harvest strategy across all environmental and evaluation contexts. No harvest strategy consistently dominated across all environmental or evaluation contexts, and each harvest strategy could be perceived as an optimal (or near optimal) choice in every environmental context (Figure 3, Supporting Information S2.2). Raw scores (Figure 3) show the challenges of achieving sustainability outcomes in harvest systems: there were few scores of 100, which represent a situation satisfying multiple objectives without compromise, exposing the system to the likelihood of different harvest strategy preferences depending on the evaluation context. Overall, threshold-increasing-proportions was an optimal (or jointly optimal) strategy in 50% of contexts, followed by proportional (44%), threshold-proportional (35%) and constant (32%; Supporting Information S2.2). However threshold-proportional had a higher average ranking (73), followed by proportional (71), then threshold-increasing-proportions and constant (64 and 39 respectively). All adaptive strategies had higher relative and raw mean scores (87-90, and 54-56 respectively) compared to constant harvest strategies (74 and 45 respectively for relative and raw scores; Supporting Information S2.2). Generally, in the cases where the biggest gains can be made by selecting the best strategy (including for composite sets of *population-focus*, *classic*, and *classic pop.+harv*., as well as for faster life history species), threshold and adaptive strategies are preferred over constant harvest. Similarly, in cases where proportional harvest performed optimally, threshold-based strategies were typically close behind, whereas there are several cases where threshold strategies outperform proportional harvest. Cases were constant harvests could perform well were often composite metric sets focussing on harvest or complete-small-game (i.e. with no concern about overpopulation), usually with Rawlsian ethics, and typically for cases with poor average raw scores (Supporting Information S2.2).

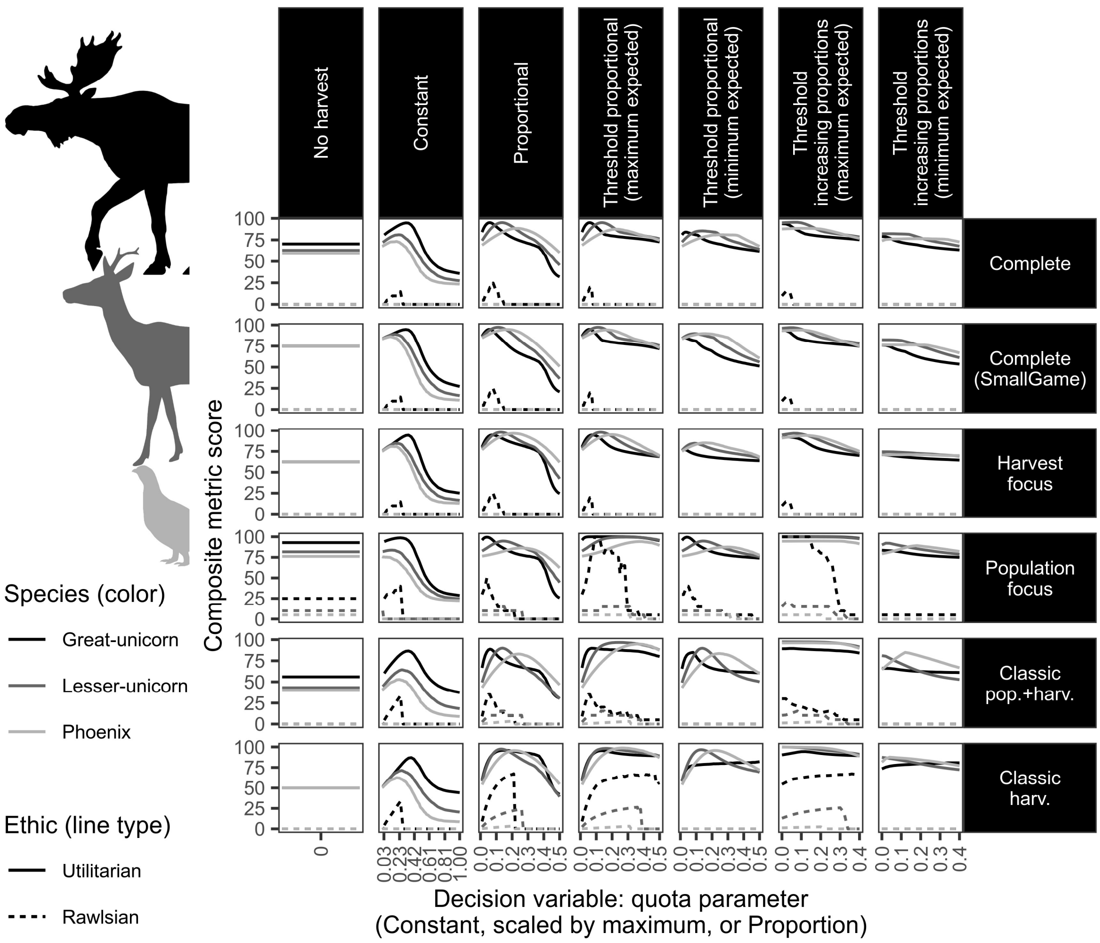
The results from our decision tree analyses emphasised the influence of choice of composite metric sets and ethics in both determining optimal strategies, and perceived sustainability (Figure 4, Supporting Information S2.3). While use of the Utilitarian ethic often produced outcomes perceived as ‘good’ (72%) or ‘relatively good’ (85%), the use of a Rawlsian ethic was not likely to produce ‘good’ outcomes (3%) unless viewed relative to other potential strategies (i.e. as ‘relatively good’; 66%; Figure 5). The latter are mainly due to outcomes being considered as equally bad. Raw Rawlsian scores were most sensitive to the addition or removal of individual metrics, but differences were also observable in Utilitarian scores (Figure 6). Often the more holistic sets (i.e. those including metrics from both harvest and population domains, alongside non-extinction) showed worse outcomes than simpler sets. These trends differed between harvest strategies: the performance gap between constant or proportional harvest strategies and the more complex harvest strategies was larger when evaluated using simplistic composite sets, relative to when more holistic composite sets are used.

All harvest strategies were more likely to produce ‘relatively good’ outcomes when viewed under a Utilitarian perspective (54-99%), and the majority of these were outright good for all but the constant harvest strategy (78-88% for adaptive harvests, 44% for constant harvest; Figure 6). Under the Rawlsian perspective the majority in all strategies were ‘relatively good’ (59-68%), however only the minority of cases were outright ‘good’ (<6%), and never under constant harvest. The probability of achieving at least a ‘relatively good’ score in the Utilitarian outcomes was significantly better when moving from a constant harvest strategy to an adaptive one (79-90%), however there were always exceptions (10-21% of cases; Supporting Information S2.4). Exceptions occurred mainly for the Great-unicorn (but existed at least once in every species), and were typically due to variability in harvest levels (Supporting Information S2.4).

Higher environmental variability resulted in lower scores, and moderate starting populations performed better than overabundant starting populations, and both substantially better than quasi-extinct starting populations, in terms of raw scores for both ethics, and relative Utilitarian scores (Figure 6). These trends reversed when viewed from a relative Rawlsian perspective, due to a reduction in maximum observed scores. There were substantial exceptions, for example 38% of pairwise comparisons showed higher scores for higher variability scenarios under the raw Utilitarian perspective, and 17% of quasi-extinction cases and 38% of overpopulation cases performed better than their equivalent moderate starting populations (Supporting Information S2.4). This highlights the importance of interactions between individual metrics that make up composite indices. In the species comparisons these interactions were also apparent: for Rawlsian raw scores there was the expected trend of slower life-history species providing more sustainability than faster life history species, yet for Utilitarian raw score comparisons the intermediate life-history species performed (slightly) better than others.

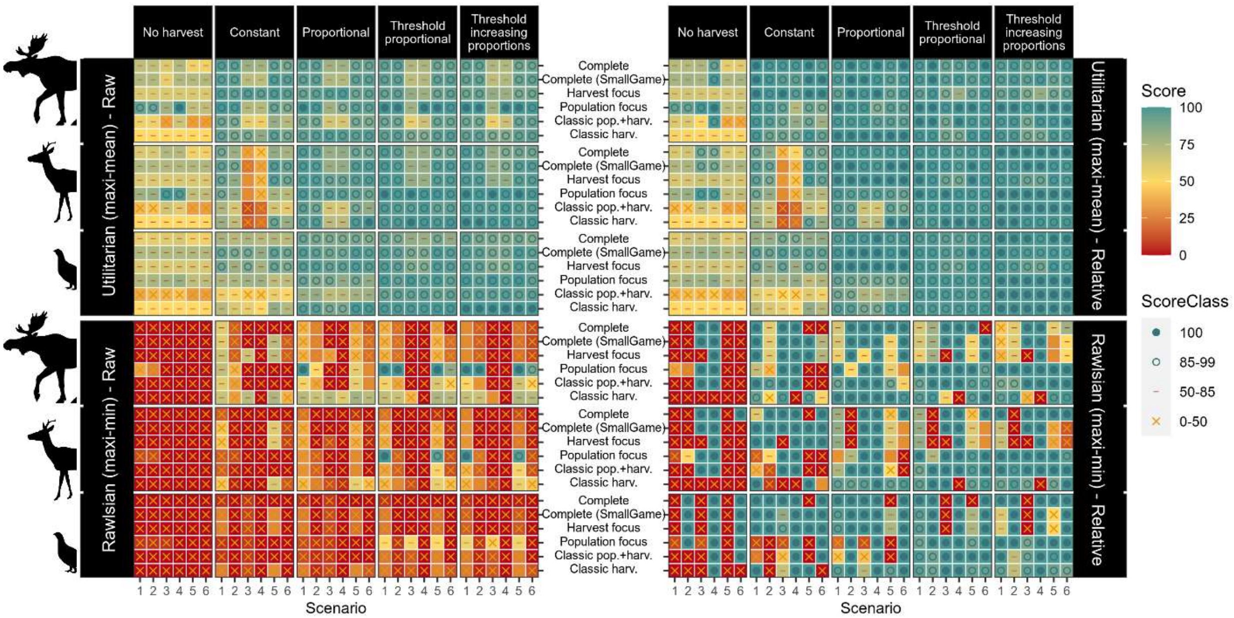
[Figure 2. Raw composite scores across quota parameters]

**Figure 2:** Raw composite scores for great-unicorn, lesser-unicorn, and phoenix, under each harvest strategy. Columns represent harvest strategies, and rows represent composite sets used for evaluation. Species are indicated by line colour, and ethic by line type. Results are for scenario 2: high variability/uncertainty and moderate starting-population sizes. Results for other scenarios, including variability, are in Supporting Information S2. For constant and proportional harvest procedures (columns two and three), the x-axis shows the constant scaled by the maximum constant, or proportion respectively. For threshold proportional and threshold increasing-proportions strategies (fourth to seventh column), the x-axis shows the proportion 1, and the score on the y-axis is expected maximum or minimum for that proportion 1 (i.e. across multiple threshold values, and gaps between proportions 1 and 2).



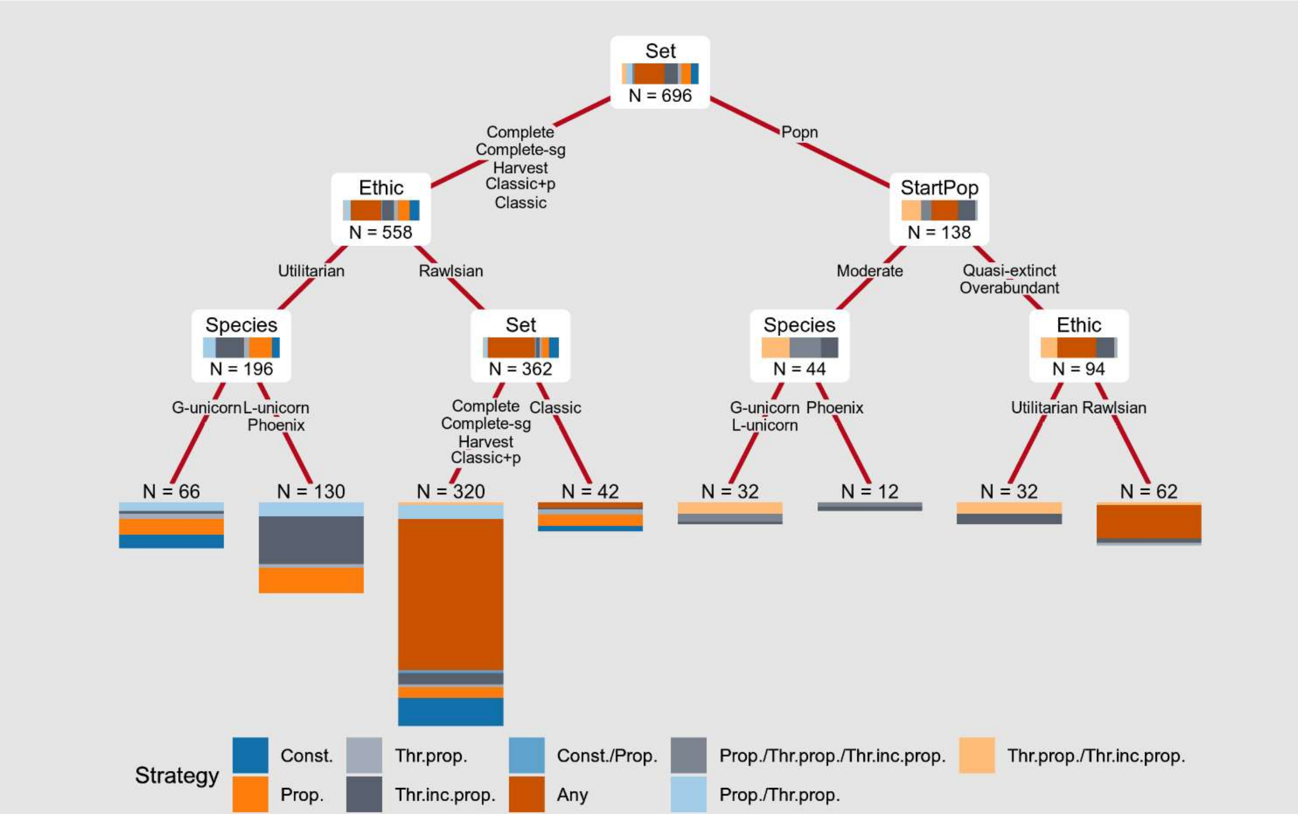
[Figure 3. Maximum composite score for each harvest strategy, across environmental and evaluation contexts]

**Figure 3**: Maximum expected composite metric scores across environmental and evaluation contexts, under a Utilitarian or Rawlsian ethic. Scores are given as raw (absolute) in the left hand panels or relative (scaled relative to other harvest strategies) in the right hand panels. Raw scores reflect the likelihood an outcome is perceived as sustainable, whereas relative scores show optimality (a strategy is 'optimal' if it receives a solid blue circle, and it is 'dominant' over all the other strategies if no other strategy also received a score of 100 for that context). Scenario codes are given in Figure 1. Alternative group summaries of optimal scores are provided in Supporting Information S2.3.



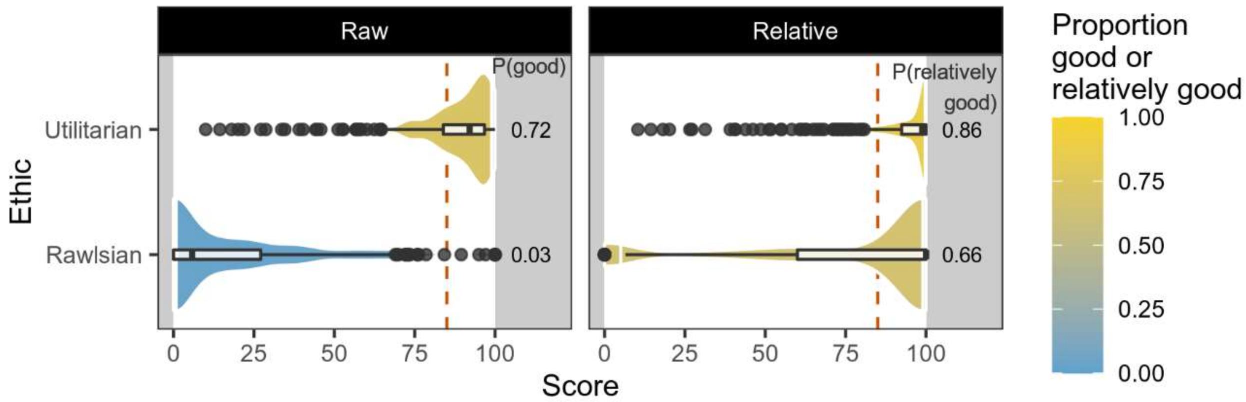
[Figure 4: Conditional inference tree for optimal strategy]

**Figure 4:** Conditional inference tree showing the choice of optimal harvest strategy (or multiple strategies if equal) under increasingly differentiated contexts. Branches diverge according to the most influential variable at that node, with branch labels indicating the distribution. Cut to depth of 3 branches for display: full tree available in Supporting Information S2.



[Figure 5: Ethical perspective comparisons]

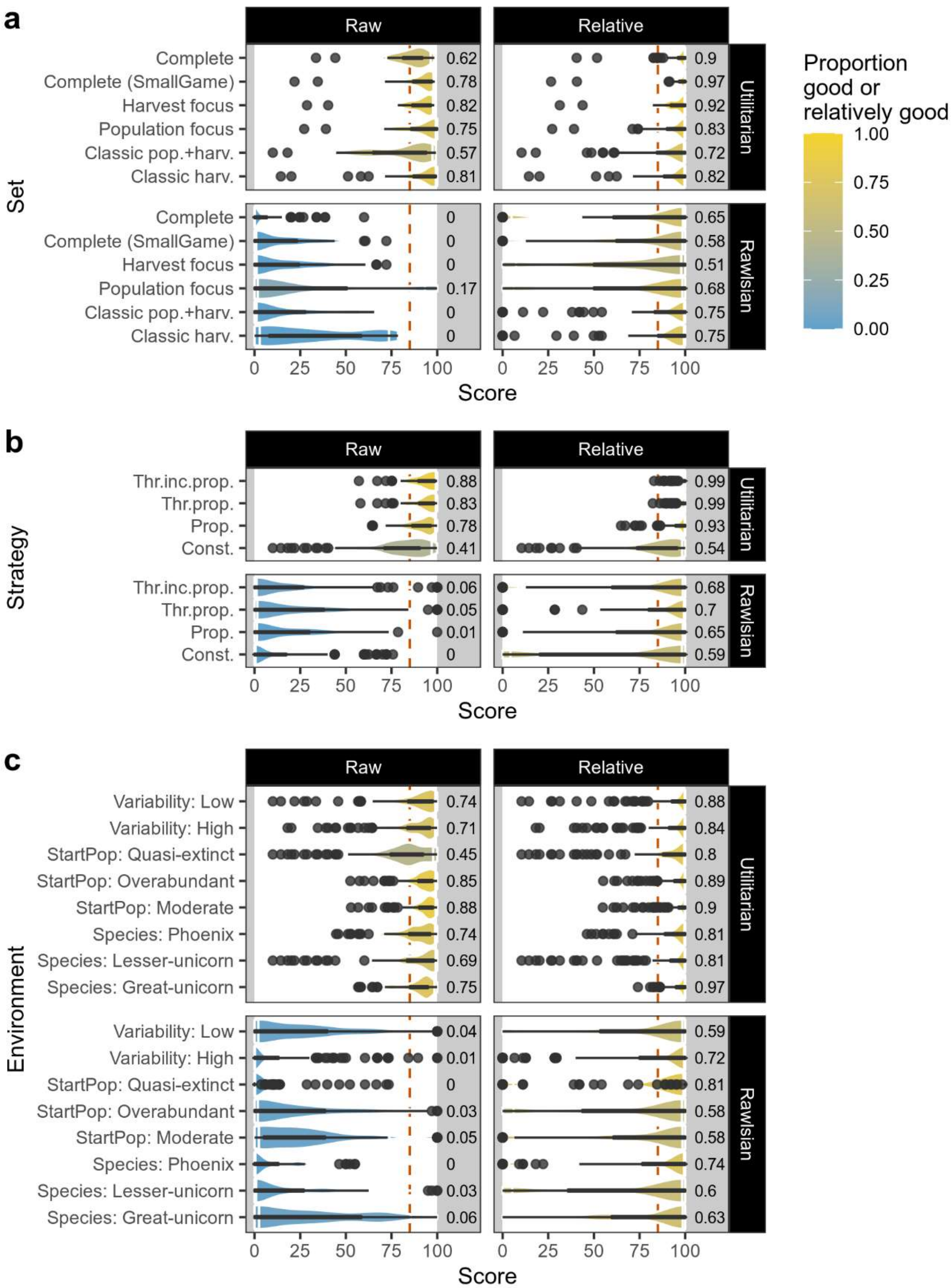
**Figure 5:** Distributions of scores by ethical perspective. Values in grey panels show proportions of ‘good’ and ‘relatively good’ for individual factors. White lines within the violin plots mark the 5% and 95% quantiles, and boxplots within the violins show median and quartiles, with whiskers extending to 1.5 times the interquartile range. For pairwise comparisons, see Supporting Information S2. There are n = 432 cases in each violin.




[Figure 6: Composite set, harvest strategy, and environmental context comparisons]

[next page]

**Figure 6:** Distributions of scores by ethical perspective. Values in grey panels show proportions of ‘good’ and ‘relatively good’ for individual factors. White lines within the violin plots mark the 5% and 95% quantiles, and boxplots within the violins show median and quartiles, with whiskers extending to 1.5 times the interquartile range. Cases in each violin: harvest strategy = 108, set comparisons = 72, environmental contexts variability = 216, species and starting population = 144. For pairwise comparisons, see Supporting Information S2.



a
b
c
Raw
Relative
Utilitarian
Rawlsian
Set
Strategy
Environment
Score
Proportion good or relatively good
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Thr.inc.prop.
Thr.prop.
Prop.
Const.
Variability: Low
Variability: High
StartPop: Quasi-extinct
StartPop: Overabundant
StartPop: Moderate
Species: Phoenix
Species: Lesser-unicorn
Species: Great-unicorn

## Discussion

Aiming to develop heuristics for sustainability in wildlife harvest systems, we ran 176,910 stochastic simulation models, and evaluated them against 12 composite sustainability indices representing different ethical perspectives and evaluation contexts. We found that no harvest strategy was optimal across all environmental and evaluation contexts tested, and every harvest strategy was at least near-optimal in at least one evaluation context in every environmental context (Figure 3). Harvest systems including thresholds or proportional harvest were more likely to deliver good outcomes, be perceived as sustainable in more varied contexts, and involved less precipitous risk of population declines compared to constant harvest, particularly when the gains possible from selecting the optimal strategy were the greatest. This supports prior analytical and review comparisons (Deroba & Bence, 2008; Engen, Lande, & Sæther, 1997; Hilker & Liz, 2020; Lande et al., 1997), and importantly, extends systematic assessment across a diversity of environmental and evaluation contexts more likely to be encountered in applied wildlife harvest management.

Dominant factors influencing sustainability of harvest systems centred around stakeholder perspectives: ethical stance, objectives considered, and whether the strategy was being assessed absolutely or relative to others (Figures 3-6, Supporting Information S2.2-2.3), highlighting the non-triviality of accounting for diverse ethical perspectives when addressing trade-offs and social equity in environmental management (Friedman et al., 2018; Law et al., 2018). In general, a Utilitarian 'aggregate good' ethic was more likely to suggest outcomes as 'good', whereas a Rawlsian 'maximise the minimum' ethic highlighted that the majority of cases have unavoidable, and often severe, trade-offs between individual stakeholder metrics. This demonstrates the inherent complexity of achieving sustainability in terrestrial wildlife harvest systems with diverse stakeholders objectives (Gren et al., 2018; Linnell et al., 2020). The dominant influence of ethic and composite set suggests that prior theoretical analyses, by focussing on maximizing harvests and limited metrics of desirable population size, present a rather narrow and potentially misleading perspective on the conflicts and sustainability of terrestrial wildlife systems in present day social settings.

Higher variability (due to stochasticity and uncertainty, including that associated with faster life-histories) was associated with reduced sustainability (Figure 6), in line with prior studies, however these trends were neither dominant nor universal. Much emphasis within the harvest literature has been on variability (stochasticity and uncertainty), typically revealing reduced sustainability with higher variability (Lande et al., 1994, 1995, 1997; Sæther et al., 1996). Our pairwise analysis showed many exceptions. In 38% of cases higher variability actually improved raw Utilitarian outcomes. Many of these exceptions are due to threshold based evaluation criteria: increased variability allows some iterations to cross desirable threshold criteria (a form of stochastic resonance; McDonnell & Abbott, 2009), without causing equivalent crossing of undesirable criteria thresholds. This result extends prior literature regarding the effectiveness of threshold-based strategies (Hilker & Liz, 2019, 2020; Lande et al., 1997) to consider impacts of threshold-based evaluation criteria. Other exceptions included 33-56% of raw Utilitarian pairwise comparisons where the intermediate life-history species, and non-ideal starting populations performed relatively well in our comparisons(Figure 6), due to reduced trade-offs between individual metrics.

Our results suggest that management of slower life-history species should be particularly concerned about low population sizes: recovery from these could be lengthy (Kritzer, Costello, Mangin, & Smith, 2019). In 'faster' species recovering from extreme low populations, harvest strategy trades off speed, magnitude, and likelihood of recovery with harvest early in the time period, a trade-off likely to depend on the productivity of the population (Babcock, McAllister, & Pikitch, 2007). Overall, this supports adaptive harvest strategies (including proportional and/or thresholds) which provide economic and ecological resilience of harvest under both scientific and environmental uncertainty, and particularly uncertainty in the face of directional threats such as climate change (Kritzer et al., 2019).

There are substantial applied management implications of trade-offs between individual metrics in different composite sets. Scores from simpler composite sets were typically higher (but not always) than more holistic sets (Figure 6, Supporting Information S2.4): perceived outcomes depended on which metrics were included, how they trade-off, and how they were combined. Two key implications can be drawn: 1) simpler 'classic' metrics commonly used in theoretical models may give a false perception of the magnitude of the benefits of more complex harvest strategies over constant harvests in some cases, and 2) the formulation of harvest objectives, particularly maximising harvests, have a strong influence in determining optimal harvest decisions. This is particularly important to consider in the context of terrestrial wildlife harvest, where there is seemingly a widespread tendency for the objective of maximizing yields to be included. It persists even in cases where extensive stakeholder and manager engagement do not indicate maximum yields as a universally valued objective, and even while recognising the strong trade-off between population stability and harvest goals (Johnson et al., 1997, 2019). In all of our simulated species the critical thresholds for management were often well below theoretical maximum sustainable yield levels (Supporting Information S1). Inclusion of yield maximization is likely due to the classic tradition of yield being the sole focus of 'sustainability' in wildlife harvest, despite development of more diverse definitions (Quinn & Collie, 2005). In fisheries contexts where yield is measured in tonnage this may be appropriate, but in contemporary, predominantly recreational terrestrial wildlife harvest there is no *a priori* reason to value maximizing mean harvests above or even equally to other objectives, especially given the diversity of human-wildlife conflicts associated with high density populations (Linnell et al. 2020).

This large potential for conflicts and trade-offs emphasises that wildlife harvest decisions are likely to benefit from tools designed for decision-making under conflict and complexity. This includes MSE models to evaluate and compare outcomes for multiple models, actions, and metrics (Bunnefeld et al., 2011; Marasco et al., 2007; Punt et al., 2016), and Structured Decision Making (SDM) tools for management of conflicts through stakeholder negotiations (Mitchell et al., 2018; Robinson et al., 2016). Avoiding exacerbating conflicts is endorsed in environmental management (Redpath et al., 2013); our analysis demonstrates how MSE can map out potential for conflict, and thereby contribute to this approach.

Given our aim of developing heuristics across a range of species contexts for a set of harvest strategies, we developed our model using a consistent but relatively simple population dynamics framework: one closed-population harvested species, undifferentiated by age, sex, or spatially, logistic growth and simple characterisations of uncertainty and variability, single decision rules being applied over the whole time frame, and no time-discounting or monetary valuation of costs and benefits. We discuss these issues as they pertain to this analysis more in the full model description in the Supporting Information S1. We also do not consider starting conditions for stakeholders (e.g. current entitlement), which can severely constrain management decisions in practice (Mitchell et al., 2018). While alternative assumptions may change the particulars of results, even the simple assumptions we employed resulted in many complex trade-offs among the diverse metrics evaluated, and we would expect the main conclusion of context dependency and importance of evaluation perspective to hold.

## Conclusions

Sustainability is a central, but often elusive goal of wildlife harvest management, challenged by complex socio-ecological systems, with many potential conflicts and uncertainties. Our stochastic simulation analysis provides the first detailed and consistent comparison of multiple sustainability metrics, across a representative range of common terrestrial wildlife harvest systems. While we conclude, similarly to prior studies, that adaptive harvest systems including thresholds and proportional harvest were more likely to be perceived as sustainable in more varied contexts compared to constant harvest, our analysis reveals the many exceptions to such heuristics. We found that the strongest driver of perceived outcomes was the evaluation framing, rather than environmental contexts. Indeed, in every environmental context all strategies could be perceived as optimal in at least one evaluation framing. Two key results for applied management are, first, that outcomes based on simplified metrics (e.g. non-extinction and maximizing mean harvest only) popular in the theoretical literature may give misleading impressions of the relative benefits of different harvest systems in applied contexts, and second, that harvest maximization has strong and potentially undue influence on analyses of 'optimality' in terrestrial wildlife harvest contexts. Our results highlight that trade-offs between sustainability objectives are largely inevitable, and, with no single optimum strategy, 'optimal' harvest systems need to be identified with careful consideration of the appropriateness of sustainability metrics and the ethical implications of their combination.

## Authors' contributions

All authors contributed to the conceptualisation of the analysis. EL developed the methodology, conducted simulations, analysed the data, and led the writing of the manuscript. All authors contributed critically to the analysis and drafts, and gave final approval for publication.

## Acknowledgements

Erling Solberg for discussions on moose management. This study was funded by the Research Council of Norway (grant 251112).

## Data availability

Input data, simulation code and results are available in OSF repository DOI 10.17605/OSF.IO/U52RP (https://osf.io/u52rp/?view_only=e36abdca3e3c45d8813e6f7b20ce159a)

Analysis code and results are available in OSF repository DOI 10.17605/OSF.IO/CGWA6 (https://osf.io/cgwa6/?view_only=973dda4c88ea4a008c3b6e58ff149822)

## References

Artelle, K. A., Reynolds, J. D., Treves, A., Walsh, J. C., Paquet, P. C., & Darimont, C. T. (2018). Hallmarks of science missing from North American wildlife management. *Science Advances*, *4*(3), eaao0167. doi: 10.1126/sciadv.aao0167

Babcock, E. A., McAllister, M. K., & Pikitch, E. K. (2007). Comparison of Harvest Control Policies for Rebuilding Overfished Populations within a Fixed Rebuilding Time Frame. *North American Journal of Fisheries Management*, *27*(4), 1326–1342. doi: 10.1577/M06-124.1

Bielby, J., Mace, G. M., Bininda-Emonds, O. R. P., Cardillo, M., Gittleman, J. L., Jones, K. E., … Purvis, A. (2007). The Fast-Slow Continuum in Mammalian Life History: An Empirical Reevaluation. *The American Naturalist*, *169*(6), 748–757. doi: 10.1086/516847

Bischof, R., Nilsen, E. B., Brøseth, H., Männil, P., Ozoliņš, J., & Linnell, J. D. C. (2012). Implementation uncertainty when using recreational hunting to manage carnivores. *Journal of Applied Ecology*, *49*(4), 824–832. doi: 10.1111/j.1365-2664.2012.02167.x

Bled, F., & Belant, J. L. (2019). demetR: a Bayesian population simulation web-application for harvest management. *Ursus*, *29*(2), 82–92. doi: 10.2192/URSUS-D-18-00012.1

Borkovec, M., Madin, N., Wickham, H., Chang, W., Henry, L., Pedersen, T. L., … Yutani, H. (2019). ggparty: “ggplot” Visualizations for the “partykit” Package (Version 1.0.0). Retrieved from https://CRAN.R-project.org/package=ggparty

Bunnefeld, N., Hoshino, E., & Milner-Gulland, E. J. (2011). Management strategy evaluation: a powerful tool for conservation? *Trends in Ecology & Evolution*, *26*(9), 441–447. doi: 10.1016/j.tree.2011.05.003

Corlatti, L., Sanz-Aguilar, A., Tavecchia, G., Gugiatti, A., & Pedrotti, L. (2019). Unravelling the sex- and age-specific impact of poaching mortality with multievent modeling. *Frontiers in Zoology*, *16*(1), 20. doi: 10.1186/s12983-019-0321-1

Davis, K. J., Chadès, I., Rhodes, J. R., & Bode, M. (2019). General rules for environmental management to prioritise social ecological systems research based on a value of information approach. *Journal of Applied Ecology*, *56*(8), 2079–2090. doi: 10.1111/1365-2664.13425

Deroba, J. J., & Bence, J. R. (2008). A review of harvest policies: Understanding relative performance of control rules. *Fisheries Research*, *94*(3), 210–223. doi: 10.1016/j.fishres.2008.01.003

DeVore, R. M., Butler, M. J., Wallace, M. C., & Liley, S. G. (2018). Population Dynamics Model to Inform Harvest Management of a Small Elk Herd in Central New Mexico. *Journal of Fish and Wildlife Management*, *9*(2), 531–544. doi: 10.3996/012018-JFWM-008

Engen, S., Lande, R., & Sæther, B.-E. (1997). Harvesting Strategies for Fluctuating Populations Based on Uncertain Population Estimates. *Journal of Theoretical Biology*, *186*(2), 201–212. doi: 10.1006/jtbi.1996.0356

Eriksen, L. F., Moa, P. F., & Nilsen, E. B. (2018). Quantifying risk of overharvest when implementation is uncertain. *Journal of Applied Ecology*, *55*(2), 482–493. doi: 10.1111/1365-2664.12992

Friedman, R. S., Law, E. A., Bennett, N. J., Ives, C. D., Thorn, J. P. R., & Wilson, K. A. (2018). How just and just how? A systematic review of social equity in conservation research. *Environmental Research Letters*, *13*(5), 053001. doi: 10.1088/1748-9326/aabcde

Gren, I.-M., Häggmark-Svensson, T., Elofsson, K., & Engelmann, M. (2018). Economics of wildlife management—an overview. *European Journal of Wildlife Research*, *64*(2), 22. doi: 10.1007/s10344-018-1180-3

Hampton, J. O., Warburton, B., & Sandøe, P. (2019). Compassionate versus consequentialist conservation. *Conservation Biology*, *33*(4), 751–759. doi: 10.1111/cobi.13249

Hayward, M. W., Callen, A., Allen, B. L., Ballard, G., Broekhuis, F., Bugir, C., … Wüster, W. (2019). Deconstructing compassionate conservation. *Conservation Biology*, *33*(4), 760–768. doi: 10.1111/cobi.13366

Hilborn, R., Fulton, E. A., Green, B. S., Hartmann, K., Tracey, S. R., & Watson, R. A. (2015). When is a fishery sustainable? *Canadian Journal of Fisheries and Aquatic Sciences*, *72*(9), 1433–1441. doi: 10.1139/cjfas-2015-0062

Hilker, F. M., & Liz, E. (2019). Proportional threshold harvesting in discrete-time population models. *Journal of Mathematical Biology*, *79*(5), 1927–1951. doi: 10.1007/s00285-019-01415-7

Hilker, F. M., & Liz, E. (2020). Threshold harvesting as a conservation or exploitation strategy in population management. *Theoretical Ecology*. doi: 10.1007/s12080-020-00465-8

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, *15*(3), 651–674. doi: 10.1198/106186006X133933

Hothorn, T., & Zeileis, A. (2015). partykit: A Modular Toolkit for Recursive Partytioning in R. *Journal of Machine Learning Research*, *16*(118), 3905–3909.

Johnson, F. A., Moore, C. T., Kendall, W. L., Dubovsky, J. A., Caithamer, D. F., Kelley, J. R., & Williams, B. K. (1997). Uncertainty and the Management of Mallard Harvests. *The Journal of Wildlife Management*, *61*(1), 202–216. JSTOR. doi: 10.2307/3802429

Johnson, F. A., Zimmerman, G. S., Huang, M. T., Padding, P. I., Balkcom, G. D., Runge, M. C., & Devers, P. K. (2019). Multi-species duck harvesting using dynamic programming and multi-criteria decision analysis. *Journal of Applied Ecology*, *56*(6), 1447–1459. doi: 10.1111/1365-2664.13377

Kameda, T., Inukai, K., Higuchi, S., Ogawa, A., Kim, H., Matsuda, T., & Sakagami, M. (2016). Rawlsian maximin rule operates as a common cognitive anchor in distributive justice and risky decisions. *Proceedings of the National Academy of Sciences*, *113*(42), 11817. doi: 10.1073/pnas.1602641113

Kappes, A., Kahane, G., & Crockett, M. J. (2016). From risk to fairness. *Proceedings of the National Academy of Sciences*, *113*(42), 11651. doi: 10.1073/pnas.1614111113

Kritzer, J. P., Costello, C., Mangin, T., & Smith, S. L. (2019). Responsive harvest control rules provide inherent resilience to adverse effects of climate change and scientific uncertainty. *ICES Journal of Marine Science*, *76*(6), 1424–1435. doi: 10.1093/icesjms/fsz038

Lacy, R. C. (1993). VORTEX: a computer simulation model for population viability analysis. *Wildlife Research*, *20*(1), 45–65. doi: 10.1071/wr9930045

Lande, R., Engen, S., & Saether, B.-E. (1994). Optimal harvesting, economic discounting and extinction risk in fluctuating populations. *Nature*, *372*(6501), 88–90. doi: 10.1038/372088a0

Lande, R., Engen, S., & Saether, B.-E. (1995). Optimal Harvesting of Fluctuating Populations with a Risk of Extinction. *The American Naturalist*, *145*(5), 728–745. JSTOR. Retrieved from JSTOR.

Lande, R., Sæther, B.-E., & Engen, S. (1997). Threshold Harvesting for Sustainability of Fluctuating Resources. *Ecology*, *78*(5), 1341–1350. JSTOR. doi: 10.2307/2266129

Law, E. A., Bennett, N. J., Ives, C. D., Friedman, R., Davis, K. J., Archibald, C., & Wilson, K. A. (2018). Equity trade-offs in conservation decision making. *Conservation Biology*, *32*(2), 294–303. doi: 10.1111/cobi.13008

Leung, B., Finnoff, D., Shogren, J. F., & Lodge, D. (2005). Managing invasive species: Rules of thumb for rapid assessment. *Ecological Economics*, *55*(1), 24–36. doi: 10.1016/j.ecolecon.2005.04.017

Linnell, J. D. C., Cretois, B., Nilsen, E. B., Rolandsen, C. M., Solberg, E. J., Veiberg, V., … Kaltenborn, B. (2020). The challenges and opportunities of coexisting with wild ungulates in the human-dominated landscapes of Europe's Anthropocene. *Biological Conservation*, *244*, 108500. doi: 10.1016/j.biocon.2020.108500

Manning, S. E., Stevens, B. S., & Williams, D. M. (2019). Simulated performance of multi-year harvest regulation cycles for wild turkeys. *The Journal of Wildlife Management*, *83*(5), 1032–1042. doi: 10.1002/jwmg.21678

Marasco, R. J., Goodman, D., Grimes, C. B., Lawson, P. W., Punt, A. E., & Quinn II, T. J. (2007). Ecosystem-based fisheries management: some practical suggestions. *Canadian Journal of Fisheries and Aquatic Sciences*, *64*(6), 928–939. doi: 10.1139/f07-062

McDonnell, M. D., & Abbott, D. (2009). What Is Stochastic Resonance? Definitions, Misconceptions, Debates, and Its Relevance to Biology. *PLOS Computational Biology*, *5*(5), e1000348. doi: 10.1371/journal.pcbi.1000348

McGowan, C. P., Runge, M. C., & Larson, M. A. (2011). Incorporating parametric uncertainty into population viability analysis models. *Biological Conservation*, *144*(5), 1400–1408. doi: 10.1016/j.biocon.2011.01.005

Mersmann, O., Trautmann, H., Steuer, D., & Bornkamp, B. (2018). truncnorm: Truncated Normal Distribution (Version 1.0-8). Retrieved from https://CRAN.R-project.org/package=truncnorm

Microsoft Corporation, M., & Weston, S. (2019). doSNOW: Foreach Parallel Adaptor for the “snow” Package (Version 1.0.18). Retrieved from https://CRAN.R-project.org/package=doSNOW

Miller, J. A. O., Furness, R. W., Trinder, M., & Matthiopoulos, J. (2019). The sensitivity of seabird populations to density-dependence, environmental stochasticity and anthropogenic mortality. *Journal of Applied Ecology*, *56*(9), 2118–2130. doi: 10.1111/1365-2664.13448

Mitchell, M. S., Cooley, H., Gude, J. A., Kolbe, J., Nowak, J. J., Proffitt, K. M., … Thompson, M. (2018). Distinguishing values from science in decision making: Setting harvest quotas for mountain lions in Montana. *Wildlife Society Bulletin*, *42*(1), 13–21. doi: 10.1002/wsb.861

Nilsen, E. B. (2017). The Use of Quantitative Models in the Harvest Management of Wild Ungulates, Carnivores and Small Game: Using Norway as a Case Study. In E. J. Milner-Gulland, E. Nicholson, & N. Bunnefeld (Eds.), *Decision-Making in Conservation and Natural Resource Management: Models for Interdisciplinary Approaches* (pp. 182–195). Cambridge: Cambridge University Press. doi: 10.1017/9781316135938.008

Pellikka, J., Kuikka, S., Lindén, H., & Varis, O. (2005). The role of game management in wildlife populations: uncertainty analysis of expert knowledge. *European Journal of Wildlife Research*, *51*(1), 48–59. doi: 10.1007/s10344-004-0073-9

Punt, A. E., Butterworth, D. S., Moor, C. L. de, Oliveira, J. A. A. D., & Haddon, M. (2016). Management strategy evaluation: best practices. *Fish and Fisheries*, *17*(2), 303–334. doi: 10.1111/faf.12104

Quinn, T. J., & Collie, J. S. (2005). Sustainability in single-species population models. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1453), 147–162. doi: 10.1098/rstb.2004.1577

R Core Team. (2020). R: A Language and Environment for Statistical Computing (Version 3.6.3). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Redpath, S. M., Young, J., Evely, A., Adams, W. M., Sutherland, W. J., Whitehouse, A., … Gutiérrez, R. J. (2013). Understanding and managing conservation conflicts. *Trends in Ecology & Evolution*, *28*(2), 100–109. doi: 10.1016/j.tree.2012.08.021

Riley, S. J., Decker, D. J., Enck, J. W., Curtis, P. D., Lauber, T. B., & Brown, T. L. (2003). Deer populations up, hunter populations down: Implications of interdependence of deer and hunter population dynamics on management. *Écoscience*, *10*(4), 455–461. JSTOR. Retrieved from JSTOR.

Robinson, K. F., Fuller, A. K., Hurst, J. E., Swift, B. L., Kirsch, A., Farquhar, J., … Siemer, W. F. (2016). Structured decision making as a framework for large-scale wildlife harvest management decisions. *Ecosphere*, *7*(12), e01613. doi: 10.1002/ecs2.1613

Sæther, B.-E., Engen, S., & Lande, R. (1996). Density-Dependence and Optimal Harvesting of Fluctuating Populations. *Oikos*, *76*(1), 40–46. JSTOR. doi: 10.2307/3545746

Stevens, B. S., Bence, J. R., Porter, W. F., & Parent, C. J. (2017). Structural uncertainty limits generality of fall harvest strategies for wild turkeys. *The Journal of Wildlife Management*, *81*(4), 617–628. doi: 10.1002/jwmg.21228

Strasser, H., & Weber, C. (1999). On the Asymptotic Theory of Permutation Statistics [Paper]. Retrieved September 16, 2020, from https://epub.wu.ac.at/102/

van Vliet, N., & Nasi, R. (2019). What do we know about the life-history traits of widely hunted tropical mammals? *Oryx*, *53*(4), 670–676. doi: 10.1017/S0030605317001545

Weinbaum, K. Z., Brashares, J. S., Golden, C. D., & Getz, W. M. (2013). Searching for sustainability: are assessments of wildlife harvests behind the times? *Ecology Letters*, *16*(1), 99–111. doi: 10.1111/ele.12008

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. doi: 10.21105/joss.01686

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., … RStudio. (2020). ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics (Version 3.3.0). Retrieved from https://CRAN.R-project.org/package=ggplot2

Wickham, H., Pedersen, T. L., & RStudio. (2019). gtable: Arrange “Grobs” in Tables (Version 0.3.0). Retrieved from https://CRAN.R-project.org/package=gtable

Wilke, C. O. (2019). cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2” (Version 1.0.0). Retrieved from https://CRAN.R-project.org/package=cowplot

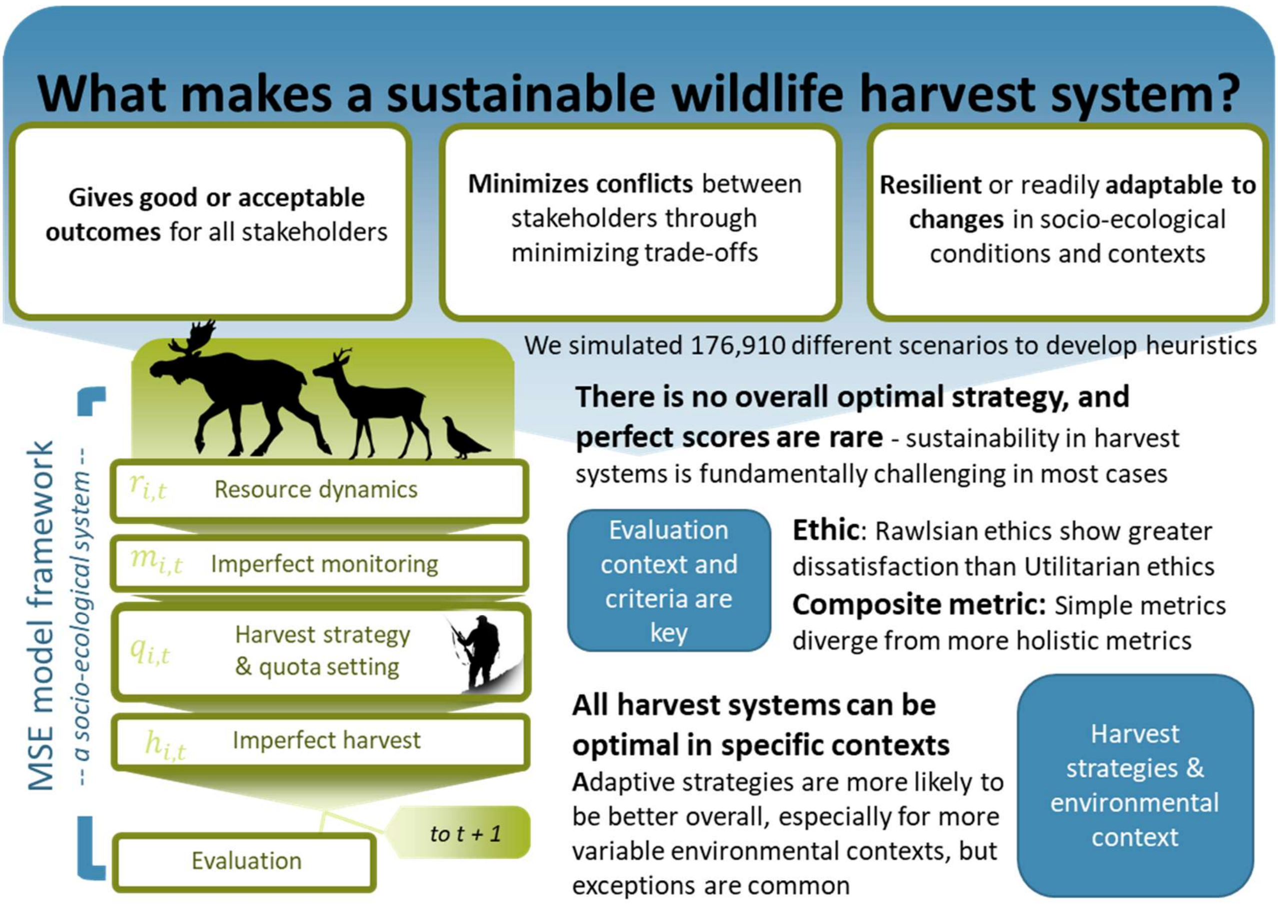
Williams, C. K. (2013). Accounting for wildlife life-history strategies when modeling stochastic density-dependent populations: A review. *The Journal of Wildlife Management*, *77*(1), 4–11. doi: 10.1002/jwmg.429

Wilson, K. A., & Law, E. A. (2016). Ethics of Conservation Triage. *Frontiers in Ecology and Evolution*, *4*. doi: 10.3389/fevo.2016.00112

## Figures and Tables

[currently placed in text for review]

## Graphical abstract



# Supporting information S1: MSE Model details

Accompanying manuscript: Sustainability of wildlife harvest in stochastic social-ecological systems.

Authors: Elizabeth Law, John D. C. Linnell, Bram van Moorter, Erlend B. Nilsen.

This supporting information repeats the methods presented in the main text, with additional detail where required, particularly regarding the assumptions and caveats.

## S1.1 Model framework

We develop an MSE model that generalises a terrestrial wildlife harvest system, with components of 1) resource dynamics, 2) monitoring observations, 3) quota setting, 4) harvest implementation, and 5) sustainability evaluation. Simulations occur in discrete yearly time steps ($t$), across a time series of 20 years (broadly considered long term for applied management plans), with multiple stochastic iterations ($i$ = 1000) per scenario. An overall perspective is provided in figure S1.1.

The **resource population component** simulates growth of a population $N_{i,t}$, via a logistic growth function determined by the population intrinsic growth rate, $r_{i,t}$, and the carrying capacity, $K$.

(1) $$N_{i,t+1} = N_{i,t} + r_{i,t} N_{i,t} \frac{(K - N_{i,t})}{K}$$ (rounded to nearest positive integer)

where:

(2) $$r_{i,t} \approx N(r_i^m, r_i^{sd})$$

(3) $$r_i^m \approx N(r^{mm}, r^{sdm})$$

(4) $$r_i^{sd} \approx N(r^{msd}, r^{sdsd})$$

We assume that $r^{mm}$, $r^{sdm}$, $r^{msd}$, $r^{sdsd}$, and $K$ are constant for each species context and variation scenario. Parameters for these are given below. Thus, variation in $r$ is simulated by a normal distribution, equally partitioned over years and iterations. Yearly variation is conceptualised to encompass the concepts of survival (due to all causes aside from hunting), reproduction, environmental variability, and demographic stochasticity; irresolvable variation, but can be low or high. Iteration variability simulates parameter uncertainty; resolvable through improved knowledge, and can be low or high. Lack of variation in $K$ assumes that the fundamental carrying capacity of the system remains the same throughout the time period assessed.

We note that the use of a standard model framework with logistic growth, applied across a fast – slow gradient of species is a simplification, as density dependent (and depensatory) effects are likely to correlate with species position along this gradient (Stevens, Bence, Porter, & Parent, 2017; Williams, 2013). A potential modification of this to better capture differences in the fast – slow species gradient might be to use a generalised theta-logistic model, with a theta < 1 for 'fast' $r$-selected species, and a theta > 1 for 'slow' $K$-selected species. However while this is conceptually practical, such parameters are challenging to estimate in application, and there is likely to be evolutionary interactions on theta with experienced environmental variability (Williams, 2013). Furthermore, Sæther et al. (1996) show that environmental stochasticity can have a larger effect on optimal harvesting strategy than the form of density-dependence. Further caveats include that we do not explicitly consider Allee effects at low population sizes (Lacy & Pollak, J.P., 2020), the full range of possible population dynamics (Saunders, Cuthbert, & Zipkin, 2018; Stevens et al., 2017; Williams, 2013). Different assumptions on the relationship between population growth rates and environmental variability are certainly possible (Colchero et al., 2019) and may induce feedbacks at a system level (Vilar & Rubi, 2018). We chose to use a simple logistic model also because of our focus on developing basic heuristics and cross-species comparisons: it provides comparability over our range of hypothetical species contexts using common, reasonable model assumptions.

This model assumes unstructured population dynamics with no spatial dynamics. As such, it ignores the impacts of age, sex, connectivity, and spatial structure in harvest systems (Colchero et al., 2019; Miller et al., 2019; Milner, Nilsen, & Andreassen, 2007). Susceptibility of different age classes to environmental variability can have significant feedbacks on population growth rates, likely to be particularly important in species where juvenile conditions correlate with adult fertility (such as the ungulates we model here) (Colchero et al 2019). Sex biases in harvest often, but not always, increase the negative impacts of hunting on a population (Milner, Nilsen, & Andreassen, 2007). We also assume that populations are closed, which can accentuate population declines, and are thus a more precautionary approach to employ (Miller, Furness, Trinder, & Matthiopoulos, 2019), at least from the perspective of population persistence.

The **monitoring component** is simulated by a single variation factor ($m_{i,t}$) acting on $N_{i,t}$, to give an estimate of the population size ($\widehat{N_{i,t}}$), to be used as the basis for management decisions.

(5) $$\widehat{N_{i,t}} = N_{i,t}\,(1 + m_{i,t})$$

where:

(6) $$m_{i,t} \approx N(m_i^m,\ m_i^{sd})$$

(7) $$m_i^m \approx N(0,\ m^{sdm})$$

(8) $$m_i^{sd} \approx N(m^{msd},\ m^{sdsd})$$

We assume that $m^{sdm}, m^{msd}, m^{sdsd}$ are constant for each species context and variation scenario, and ultimately monitoring variation has no systematic bias overall. Monitoring variation is conceptualised to encompass all the processes of sampling and observation, monitoring data analyses, and belief formation. Variation across years simulates inaccuracy or imprecision in monitoring; potentially resolvable with improved effort or monitoring technique, and can be low or high. Variation of mean bias over replications simulates parameter uncertainty regarding bias in monitoring; resolvable with improved knowledge regarding the monitoring methodology, and can be low or high. In reality, monitoring effectiveness is likely to vary with respect to the population size: with larger populations, monitoring is likely to miss or double count more individuals, and counts potentially rounded. However, we do not consider that monitoring small populations might result in larger proportional errors.

The **management decisions component** is partitioned into two parts. First, a harvest strategy is applied, converting $\widehat{N_{i,t}}$ into an initial quota, $Q_{i,t}$, given a set of quota parameters (constants, *C1*, thresholds *T1,T2,* and proportions, *P1, P2*).

(9) $$Q_{i,t} = \begin{cases} C_1, & \widehat{N_{i,t}} \leq T_1 \\ \widehat{N_{i,t}}\left(P_1 + (P_2 - P_1)\left(\frac{\widehat{N_{i,t}} - T_2}{T_2 - T_1}\right)\right), & T_1 \leq \widehat{N_{i,t}} < T_2 \\ P_2, & T_2 \leq \widehat{N_{i,t}} \end{cases}$$

Using this definition, we construct harvest strategies defined for constant, proportional, threshold proportional, with harvest proportions either stable or increasing with population size (see parameter sets in Table S1.2). We assume that the harvest strategies and the associated parameters remain consistent through the timeframe. This equation simulates evidence-based scientific recommendations of quota size (and is therefore not rounded to an integer at this stage).

$Q_{i,t}$ is then subject to random variation ($q_{i,t}$) to simulate the political interventions that often enter the quota setting process, to give a modified quota $Q'_{i,t}$.

(10) $$Q'_{i,t} = Q_{i,t}\,(1 + q_{i,t}) \qquad \text{(rounded to nearest positive integer)}$$

(11) $$q_{i,t} \approx N(0,\ q_i^{sd})$$

Variability in the quota is designed to simulate the impacts of political processes on quota development, and can either not exist (management exactly follows scientific evidence) or can introduce a ‘high’ level of variability. We assume there is no parameter uncertainty in this case, and only allow variation over years (not iterations), and we assume no overall systematic bias in quota variation.

The **harvest implementation component** simulates imperfect harvest implementation, effected as a proportional variation ($h_{i,t}$) around $Q'_{i,t}$ to give the harvest ($H_{i,t}$). This amount is then removed from $N_{i,t}$.:

(12) $H_{i,t} = Q'_{i,t}\,(1 + h_{i,t})$ (rounded to nearest positive integer)

(13) $h_{i,t} \approx N(h_i^m,\ h_i^{sd})$

(14) $h_i^m \approx N(0,\ h^{sdm})$

(15) $h_i^{sd} \approx N(h^{msd},\ h^{sdsd})$

(16) $N_{i,t+1} = N_{i,t} - H_{i,t}$

Variation across years simulates stochasticity in the harvest, and can be low or high. This variation can be conceptualised as both environmental stochasticity (irreducible) and user-driven imperfections (reducible, for example through increased enforcement or other incentive to achieve the quota), and therefore partly reducible overall. Variation across replications simulates parameter uncertainty in regards to the bias in harvest relative to the quota; resolvable through increased knowledge of the harvesters, and trust of the harvesters in the quota, and can be low or high. We simplify this by assuming that there is an unbiased estimate of how much will be harvested given a quota. In reality, hunting efficiency may vary with respect to the quota. For example, in both moose (Hunt, 2013) and ptarmigan (Eriksen, Moa, & Nilsen, 2018) hunting effectiveness increases at low tag numbers. Our formulation of the harvest imperfection as a coefficient of variance factor means that the variance will be smaller at smaller quota sizes.

A common conceptualisation of the functionality of a quota is to limit potential ‘tragedy of the commons’ by enforcing a limit on harvest, and this might be expected to produce a bias on harvest implementation such that it is more common for harvests to be below the quota than above. However, we note that in reality, quotas are often set in systems where the ‘total allowable catch’ or the maximum possible harvest legally possible under the set quota is much higher than the intended harvest (Bischof et al 2012). This is particularly common, for example, when quotas are specified with spatial or temporal specifications, or in terms of amounts per person. This means that decision makers need to estimate the relationship between the quota and the levels of harvest they intend to be taken (Moa, Eriksen, & Nilsen, 2017). Despite these critical assumptions, there are relatively few studies on imperfect harvest implementation in terrestrial wildlife systems (Bischof et al., 2012; Eriksen et al., 2018). In the current study, the quota and harvest components together define a system in which the quota is implemented with no systematic harvest bias. We conceptualize this quota as the ‘intended harvest’, or the amount expected to be harvested. We suggest that therefore the assumption that the simulated harvest may be normally distributed around the ‘intended harvest’ is reasonable, with the caveat it is unlikely to hold in all contexts.

We do not account for feedbacks and directional bias likely in harvest implementation (Eriksen et al., 2018; Hunt, 2013), and more generally through the harvest system (Bieg, McCann, & Fryxell, 2017; Fryxell, Packer, McCann, Solberg, & Sæther, 2010).

The **evaluation component** occurs after each simulation is complete, calculating the performance of each replicate over the entire timeframe, and summarising (means, medians and quantiles) over the scenario run. Evaluation metrics are designed to reflect different potential objectives and stakeholder concerns, and cover a number of socio-ecological (i.e. population-based) and harvest-based sustainability objectives (Table 1).

As multiple sustainability metrics may be relevant to a context, we develop a number of composite metrics illustrating potentially common sets of metrics (Table 2). To simulate further different stakeholder perspectives, these composite sets may be summarised under several different ethics, and we compare two of these: Utilitarian (‘aggregate good; translatable as a sum or average of the set of metrics), and Rawlsian (‘maximin’; maximising the smallest benefit across the set of metrics). To ensure scores are comparable across harvest systems for each scenario, we scale individual metrics relative to the largest mean score achieved by any harvest system and parameter set across the environmental context. Composite metrics are compiled prior to summarization, due to non-independence of the individual metric scores (see figure S1.2).

In summary, we assume that there can be yearly variation in *r, m, q,* and *h,* and variation over replications for *r, m,* and *h.* We assume variability in *r, m,* and *h,* can be low or high, simulating partial resolvability of these phenomena. We assume variability in *q* can be zero or high. In all the variable parameters, we assume normal distributions (as specified in the above equations, using the species specific parameters given in section S1.2), with no correlation of error. We select normal distributions as we assume the variability is due to a number of different sources, and the central limit theorem would suggest that these might coalesce to a normal distribution. In the monitoring, quota, and harvest variation, for simplicity we assume a coefficient of variation function proportional to the population or quota.

We made no attempt to value the monetary aspects of harvest systems (Gren et al., 2018), nor implementation costs (Kritzer et al., 2019). We note that in our analysis, by limiting the time series to 20 years but otherwise not adjusting for time discounting, we effectively default to a zero discount until year 20, and full discount thereafter. The consequences of this are that maximizing harvest objectives can drive populations to undesirably low levels when not checked by other objectives or inherent risk of variability in the species context. This results in some scenarios – particularly noticeable in the constant harvest strategy for the ‘slow’ species – resulting in the objective of maximizing harvest (without any other constraints) causing a draw down on the population to the point at which harvests are limited by the population size (often extinction). While this is not an acceptable scenario in any definition of ‘sustainability’ we use it as a cautionary note as to what focus on certain metrics and ignorance of others may cause. Applying time discounting across the time series is likely to further increase this (undesirable) effect, even with infinite time horizons (Lande et al., 1994), as they would place more value on larger harvests in the earlier points in the time frame, and discount smaller harvests caused by population decline in later years. There is no universally applicable method for defining appropriate discount rates for non-monetary values (Botzen & van den Bergh, 2014), but here we note that despite the absence of a time discounting procedure, the limited time frame of assessment effects this phenomenon in this case.

As we are not searching for equilibria, we do not apply a ‘burn-in’ time period, but rather start the population from the initial population given by the species parameters. We also do not consider time lags in management decisions, which can be common particularly in low-knowledge scenarios (Manning, Stevens, & Williams, 2019).

We also assume no other temporal feedbacks aside from those effected by density dependence and application of the harvest strategy to generate the initial quota. However, these may be common features of management systems, for example populations may be more prone to environmental variability under high population densities, and harvesters and managers may react systematically to different population densities and quotas (Bieg, McCann, & Fryxell, 2017; Fryxell, Packer, McCann, Solberg, & Sæther, 2010).

Figure S1.1: Overall MSE model framework

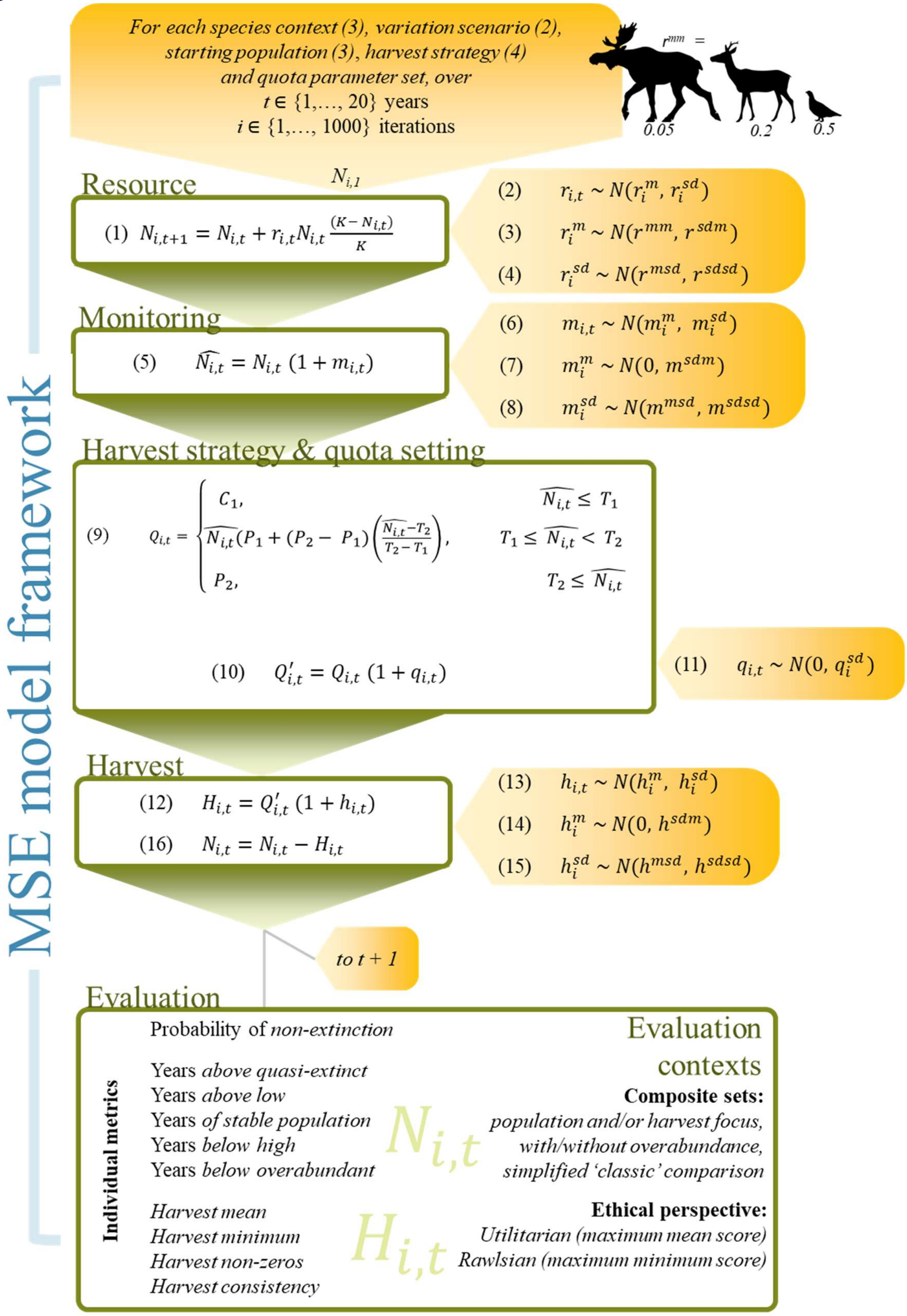
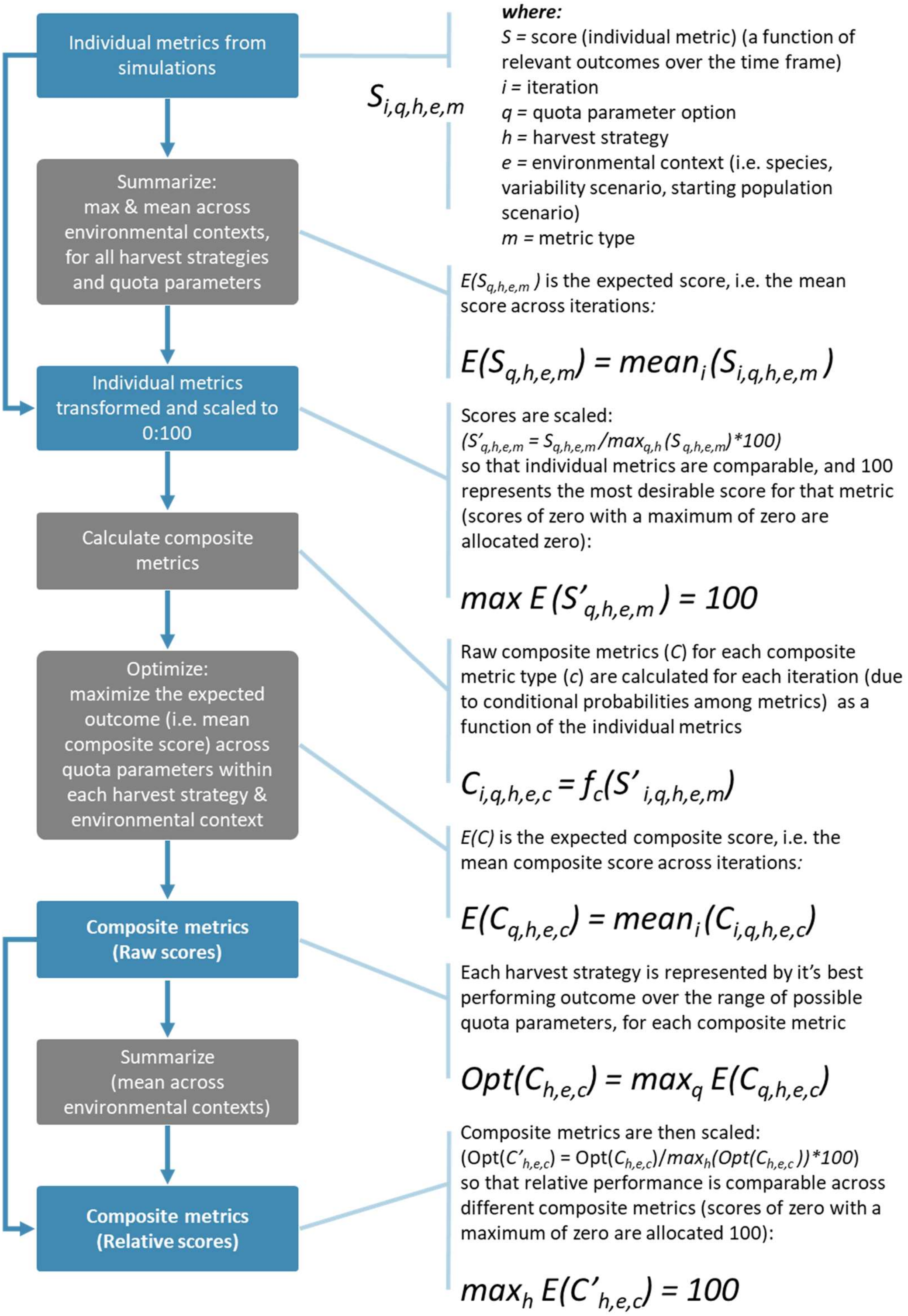


Figure S1.2: Evaluation framework: scaling and calculation of composite metrics.



## S1.2 Hypothetical species and parameters

We develop cases based on three hypothetical species spanning a range of common game species: the great-unicorn, the lesser-unicorn, and the phoenix (Table S1.1). We loosely base these hypothetical species on wildlife species harvested in a Norwegian context. To provide consistency between species, overall variation in the growth rate is specified to be equal to the species growth rate in the high variability scenario, and half of the species growth rate in the low variation scenario. For each variable parameter (here generalised to $x$), total variation ($x^{TSD}$) is split between replications and years, by partitioning the overall standard deviation equally into the iteration level standard deviation (that determines the vector of the parameter over the years, $x^{sdm}$), and the scenario level mean standard deviation (that determines the parameter mean value over the iterations, $x^{msd}$).

(17) $$x^{msd} = x^{sdm} = \frac{x^{TSD}}{2}$$

The standard deviation of the iteration level standard deviation ($x^{sdsd}$)was defined at 1/3 of the standard deviation of the mean standard deviation. Simulations of $x_i^{sd}$ were truncated to remain positive, at a minimum of 0.0001.

(18) $$x^{sdsd} = \frac{x^{sdm}}{3}$$

The great-unicorn resembles a large ungulate (e.g. moose, *Alces alces*). It is assumed to have a relatively low growth rate, carrying capacity, monitoring variation, and critical thresholds (Table S1.1). A description of moose population and harvest dynamics in a Scandinavian context is available in Sæther et al. (2001).

The lesser-unicorn resembles a small ungulate (e.g. roe deer, *Capreolus capreolus*), with a moderate growth rate, carrying capacity, monitoring variation, and critical thresholds. A description of roe deer population and harvest dynamics is available in Andersen et al. (1998).

The phoenix is reflective of a game bird (e.g. willow ptarmigan, *Lagopus lagopus*), with a relatively large potential growth rate, carrying capacity, monitoring variation, and critical thresholds. A description of willow ptarmigan population and harvest dynamics in a Scandinavian context is available in Eriksen et al. (Eriksen et al., 2018).

From these parameters, we can calculate the standard maximum sustainable yield (*MSY*) conditions given no stochasticity, occurring at $K/2$, and with an annual harvest of $rK/4$ under the logistic growth assumption (with values rounded to the nearest integer). For the great-unicorn, MSY is expected at a population of 1112 (notably larger than the moderate starting population, the high critical threshold, and close to the overabundant critical threshold), allowing an annual harvest of 28 individuals. For the lesser-unicorn, MSY is expected at a population size of 13900 (also larger than the moderate starting population, and the high critical threshold), allowing a harvest of 1390 individuals. For the phoenix, MSY is expected at 30000 (also larger than the moderate starting population, and the high critical threshold) with a harvest of 7500. We provide these calculations for comparison only: MSY using these calculations is a theoretical construct under strict assumptions and will often overestimate the true maximum sustainable yield (Quinn & Collie, 2005). We also note that, given the parameters used, the MSY population level is often higher than desirable for other stakeholder concerns, particularly in the ungulate systems.

## Table S1.1: Species parameters and variable parameter assumptions

We defined three species contexts, which specified the value of fixed constants for mean *r, K,* critical thresholds, and starting populations, and the level of variations deemed low and high for *r* and *m.* While these are loosely based on real species, the values are specified to facilitate scenario comparisons. We also provide here parameters used for quota and harvest variability, assumed to be equal across the species gradient. Variable parameters are given as the overall mean ($x^{mm}$) and overall standard deviation ($x^{TSD}$) for low and high variation scenarios; a description of how these are partitioned into yearly and iteration level distribution parameters is provided in section S1.2.

| **Component** | **Parameter – variation scenario** | | **Great-unicorn** | | **Lesser-unicorn** | | **Phoenix** | |
|---|---|---|---|---|---|---|---|---|
| | | | ***Mean*** $x^{mm}$ | ***SD*** $x^{TSD}$ | ***Mean*** $x^{mm}$ | ***SD*** $x^{TSD}$ | ***Mean*** $x^{mm}$ | ***SD*** $x^{TSD}$ |
| Resource | *r* | low<br>high | 0.05 | 0.025<br>0. 05 | 0.2 | 0.1<br>0.2 | 0.5 | 0.25<br>0.5 |
| | *K* | | 2225 | | 27800 | | 60000 | |
| Monitoring | *m* | low<br>high | 0 | 0.05<br>0.15 | 0 | 0.1<br>0.3 | 0 | 0.15<br>0.45 |
| Quota | *q* | none<br>high | 0 | 0<br>0.1 | 0 | 0<br>0.1 | 0 | 0<br>0.1 |
| Harvest | *h* | low<br>high | 0 | 0.05<br>0.25 | 0 | 0.05<br>0.25 | 0 | 0.05<br>0.25 |
| Evaluation critical thresholds | *Extinction*<br>*Quasi-extinction*<br>*Low*<br>*High*<br>*Overabundant* | | 1<br>60<br>300<br>900<br>1200 | | 1<br>167<br>2780<br>11120<br>19460 | | 1<br>5000<br>10000<br>25000<br>40000 | |
| Starting populations (respective to critical thresholds) | *Moderate start*<br>*Quasi-extinct start*<br>*Overabundant start* | | 600<br>60<br>1200 | | 6950<br>167<br>19460 | | 17500<br>5000<br>40000 | |

## S1.3 Harvest strategies, quota parameters and optimization

Harvest strategies analysed include '*constant*' (a set number of individuals harvested yearly), '*proportional*' (a set proportion of the population harvested yearly), '*threshold proportional'* (a set proportion taken yearly, provided the population is above a certain threshold), and '*threshold increasing proportions*' (provided the population is above a certain threshold, the proportion taken increases as the population size increases). These harvest strategies are defined by the quota parameters that define constants, thresholds, and proportions (Table S1.2).

Harvest strategies (also known as harvest control rules) can be either strictly followed to develop quotas, or form the principles behind quota setting (Kvamsdal et al., 2016). Here we assume the former (through eq. 9) although allow some flexibility for adjustment (through eq. 10). These harvest strategies are variably termed in the literature. Some examples:

- *Constant*: fixed-quota, constant catch (Deroba & Bence, 2008).
- *Proportional*: constant mortality rate, constant-*F*, this is one of the most commonly used rules in fisheries, often suggested to be optimal with perfect information (Deroba & Bence, 2008).
- *Threshold-proportional*: proportional threshold, developed specifically for stochastic & uncertain contexts (Engen, Lande, & Sæther, 1997); also called 'threshold' by some fisheries sources (Deroba & Bence, 2008). Note in this case, we apply the proportion with respect to the whole population, if above a threshold (c.f. applying it to the proportion of the population above the threshold).

- *Threshold-increasing-proportional*: increasing rates above a threshold, biomass-based or adjustable rate rules (Deroba & Bence, 2008). Similarly to threshold proportional, we apply the proportion with respect to the whole population, if above the threshold.

Other rules not examined here include other variations on threshold-based rules. Including constant escapement (100% take above a threshold), decreasing rates below a threshold, conditional constant catch (constant amount, unless removing that amount would exceed some predetermined maximum mortality rate) with variations on this including no take below the threshold, proportional take below the threshold. The intentions for the various rules including those not utilised here are summarised in (Deroba & Bence, 2008). Of note, the harvest strategies we analyse here are focused on the population dynamics within a system, as we do not consider the relative monetary costs and benefits of harvesting. Further harvest strategies including the monetary economics of harvesting are possible (Kvamsdal et al., 2016).

Harvest rules implemented for small game birds (grouse species) in Europe and North America are reviewed in (Moa et al., 2017). They note that proportional and threshold-proportional principles are common, however in practice bag sizes are often relatively more limited at large population sizes, against recommendations (Moa et al., 2017).

Each harvest strategy can be utilised with different quota parameters, and it is this combination (of harvest strategy and quota parameters) that forms the main ‘decision variables’ in the MSE model. To sample possible quota parameter options, we employed either a stopping rule or a grid search method, incrementally varying the parameters across the option space (Table S1.2). This search method does not cover the entire option space defined, but represents a pragmatic approach towards illustrating trade-offs across the parameter space, and optimization in relevant parameter space given the volume of parameter options available, and given the likelihood of multiple optima. While this might result in fine details of the comparisons being inaccurate, we expect the main conclusions to hold, as we saw no severe gaps in the trends across the parameter space (see Main text and Supporting Information S2).

### Table S1.2: Harvest strategies, quota parameters, and heuristics for searching the option space.

Initial harvest quotas are developed such that (as defined in eq. 9) the constant, *C1*, applies from a population of 0 until the threshold *T1*. The proportion *P1* then applies, linearly transitioning to *P2* at the threshold *T2*. After this threshold, the proportion continues at *P2*. With this same set of equations, we can define the *constant, proportional, threshold-proportional,* and *threshold-increasing-proportional* harvest strategies. We simulate over a range (option space) of quota parameters for each harvest strategy, using the increments and stopping rule (or searching the full option space).

| **Harvest procedure** | **Quota parameters (option space)** | | | | |
|---|---|---|---|---|---|
| | **C1** | **P1** | **P2** | **T1** | **T2** |
| *Constant* | 0 : stop | 0 | 0 | Inf | Inf |
| *Proportional* | 0 | 0.01 : 0.50 | = P1 | 0 | Inf |
| *Threshold-proportional* | 0 | 0.01 : 0.50 | = P1 | quasi-extinction : moderate starting population | Inf |
| *Threshold-increasing-proportional* | 0 | 0 : 0.50 | 0.1 : 1 *subject to:* P2 at 0.1-0.5 above P1 | quasi-extinction : moderate starting population | overabundant critical threshold |

| | **Increments** | **Stopping rule** |
|---|---|---|
| *Constant* | Increase C1 in increments of 1% of moderate staring population | Stop if probability of non-extinction = 0 |
| *Proportional* | Increase P1 in increments of 0.01 | Stop if probability of non-extinction = 0 |
| *Threshold-proportional* | T1 increments of 1% of (moderate starting population – quasi-extinction)<br>P1 increments of 0.01 | Full grid search |
| *Threshold-increasing-proportional* | From P1 = 0 to 0.26:<br>• T1 increments of 2.5% of (moderate starting population – quasi-extinction)<br>• P1 increments of 0.01<br>From P1 = 0.28 to 0.58:<br>• T1 increments of 5% of (moderate starting population – quasi-extinction)<br>• P1 increments of 0.02<br>Over all of these, increment the difference between P1 and P2 (the *P2 gap*) by 0.1 | Full grid search |

## S1.4 Scenarios and comparisons

To examine the performance of the harvest strategies, we first focused on comparing results for each species context and harvest strategy for scenarios where all variability in *r, m, q*, and *h* were either all low (*low variability*) or all high (*high variability*), and starting populations were at the midpoint of low and high critical thresholds (*moderate starting population*). This means that the magnitude of uncertainty was correlated between the components, however the pattern of uncertainty across years was random for all components. We then repeated the simulations with populations starting at quasi extinction (*low starting population*), and populations starting at overabundance (*high starting population*), to examine the robustness of the harvest strategies to extreme perturbations in population size and the recovery potential in such cases. Such simulations are also relevant for special management cases, for example harvest of an overabundant invasive species, or recovery of endangered species into harvestable populations. Simulations were run such that each harvest strategy and quota parameter variation is run with exactly the same starting and variable conditions (*r, m, q*, *h*) under each respective scenario (species context, variability level, and starting population size) combination.

We tested both outcomes based on the 'true' simulated populations $N_{i,t}$, as well as metrics based on the simulated monitoring data ($\widehat{N_{i,t}}$), but as the latter were virtually identical to the former in this case (as might be expected with normal distributions on errors) we report only on $N_{i,t}$.

In this analysis, we focus on the implications of alternative harvest strategies and sustainable metrics, and therefore only test the cases of 'low' and 'high' variability for each hypothetical species, and do not resolve here which sources of variability or uncertainty are most influential or valuable to address (Canessa et al., 2015; Davis, Chadès, Rhodes, & Bode, 2019).

Scenarios 3 and 4 simulate recovery of populations from quasi-extinction levels, at *low* and *high variability* scenarios respectively, via the use of a single harvest strategy set quota parameters across the entire time-frame. The great-unicorn was largely unable to reach a *stable population* level, even with zero harvest, for any variability scenario. This is not unexpected given the mean population growth rate specified for this species (r = 0.05) under the given time frame (20 years), and the critical thresholds specified (from a starting point of quasi-extinction = 60, population growth without harvest would be expected to increase the population to 159 ($60 \times (1 + 0.05)^{20}$ ), which remains below the low critical threshold = 300). There could be a higher level of recovery of great-unicorn above the *quasi-extinction* critical threshold for the more complex harvest strategies, and interestingly the high variation scenario performed considerably better than the low variation scenario in this regard, because more iterations received higher population growth rates, while thresholds minimized losses. The lesser-unicorn showed more recovery, holding *stable population* for around four years, across all harvest strategies and variation scenarios. The *stable population* outcomes were quite variable, however, and while the high variability scenario achieved all years at above the *quasi-extinction* critical threshold, the higher variability scenario achieved suboptimal scores with high levels of variability. The phoenix, with a much higher rate of population growth on average, would be expected to attain a better recovery, and had *stable population* scores only slightly lower than the baseline moderate start population scenario. The lower constant harvest rate needed to effect this increased the *above quasi-extinct* scores for this harvest strategy, but with a corresponding likely decline in *below high*.

Scenarios 5 and 6 simulate harvesting of populations starting at overabundant levels, at *low* and *high variability* scenarios respectively, via the use of a single harvest strategy set quota parameters across the entire time-frame. Results were similar to the baseline scenarios, albeit with higher *harvest mean* and lower *stable population* scores, particularly for the slower-larger species with the effect declining for phoenix. Typically differences manifested in compatibility sets being classified as high populations rather than stable populations, in both the unicorn species.

## Table S1.3: Scenarios, applied for each species and harvest procedure

| **Scenario** | **Scenario ID (SID)** | **Starting population size ($N_0$)** | ***Iteration level variability in mean:*** | | | ***Yearly variation in:*** | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | **Reproductive rate ($r$)** | **Monitoring ($m$)** | **Harvest ($h$)** | **Reproductive rate ($r$)** | **Monitoring ($m$)** | **Quota ($q$)** | **Harvest ($h$)** |
| Low variation – moderate starting population | 1 | mid | low | low | low | low | low | none | low |
| High variation – moderate starting population | 2 | mid | high | high | high | high | high | high | high |
| Low variation – low starting population | 3 | low | low | low | low | low | low | none | low |
| High variation – low starting population | 4 | low | high | high | high | high | high | high | high |
| Low variation – high starting population | 5 | high | low | low | low | low | low | none | low |
| High variation – high starting population | 6 | high | high | high | high | high | high | high | high |

## S1.5 References (for Supporting Information 1)

Andersen, R., Duncan, P., & Linnell, J. D. C. (1998). *The European roe deer: the biology of success*. Scandinavian University Press.

Bieg, C., McCann, K. S., & Fryxell, J. M. (2017). The dynamical implications of human behaviour on a social-ecological harvesting model. *Theoretical Ecology*, *10*(3), 341–354. doi: 10.1007/s12080-017-0334-3

Bischof, R., Nilsen, E. B., Brøseth, H., Männil, P., Ozoliņš, J., & Linnell, J. D. C. (2012). Implementation uncertainty when using recreational hunting to manage carnivores. *Journal of Applied Ecology*, *49*(4), 824–832. doi: 10.1111/j.1365-2664.2012.02167.x

Botzen, W. J. W., & van den Bergh, J. C. J. M. (2014). Specifications of Social Welfare in Economic Studies of Climate Policy: Overview of Criteria and Related Policy Insights. *Environmental and Resource Economics*, *58*(1), 1–33. doi: 10.1007/s10640-013-9738-8

Canessa, S., Guillera-Arroita, G., Lahoz-Monfort, J. J., Southwell, D. M., Armstrong, D. P., Chadès, I., … Converse, S. J. (2015). When do we need more data? A primer on calculating the value of information for applied ecologists. *Methods in Ecology and Evolution*, *6*(10), 1219–1228. doi: 10.1111/2041-210X.12423

Davis, K. J., Chadès, I., Rhodes, J. R., & Bode, M. (2019). General rules for environmental management to prioritise social ecological systems research based on a value of information approach. *Journal of Applied Ecology*, *56*(8), 2079–2090. doi: 10.1111/1365-2664.13425

Deroba, J. J., & Bence, J. R. (2008). A review of harvest policies: Understanding relative performance of control rules. *Fisheries Research*, *94*(3), 210–223. doi: 10.1016/j.fishres.2008.01.003

Engen, S., Lande, R., & Sæther, B.-E. (1997). Harvesting Strategies for Fluctuating Populations Based on Uncertain Population Estimates. *Journal of Theoretical Biology*, *186*(2), 201–212. doi: 10.1006/jtbi.1996.0356

Eriksen, L. F., Moa, P. F., & Nilsen, E. B. (2018). Quantifying risk of overharvest when implementation is uncertain. *Journal of Applied Ecology*, *55*(2), 482–493. doi: 10.1111/1365-2664.12992

Fryxell, J. M., Packer, C., McCann, K., Solberg, E. J., & Sæther, B.-E. (2010). Resource Management Cycles and the Sustainability of Harvested Wildlife Populations. *Science*, *328*(5980), 903–906. doi: 10.1126/science.1185802

Hunt, L. M. (2013). Using human-dimensions research to reduce implementation uncertainty for wildlife management: a case of moose (Alces alces) hunting in northern Ontario, Canada. *Wildlife Research*, *40*(1), 61–69. doi: 10.1071/WR12185

Kvamsdal, S. F., Eide, A., Ekerhovd, N.-A., Enberg, K., Gudmundsdottir, A., Hoel, A. H., … Vestergaard, N. (2016). Harvest control rules in modern fisheries management. *Elem Sci Anth*, *4*(0), 000114. doi: 10.12952/journal.elementa.000114

Manning, S. E., Stevens, B. S., & Williams, D. M. (2019). Simulated performance of multi-year harvest regulation cycles for wild turkeys. *The Journal of Wildlife Management*, *83*(5), 1032–1042. doi: 10.1002/jwmg.21678

Miller, J. A. O., Furness, R. W., Trinder, M., & Matthiopoulos, J. (2019). The sensitivity of seabird populations to density-dependence, environmental stochasticity and anthropogenic mortality. *Journal of Applied Ecology*, *56*(9), 2118–2130. doi: 10.1111/1365-2664.13448

Milner, J. M., Nilsen, E. B., & Andreassen, H. P. (2007). Demographic side effects of selective hunting in ungulates and carnivores. *Conservation Biology: The Journal of the Society for Conservation Biology*, *21*(1), 36–47. doi: 10.1111/j.1523-1739.2006.00591.x

Moa, P. F., Eriksen, L. F., & Nilsen, E. B. (2017). Harvest Regulations and Implementation Uncertainty in Small Game Harvest Management. *Frontiers in Ecology and Evolution*, *5*. doi: 10.3389/fevo.2017.00118

Quinn, T. J., & Collie, J. S. (2005). Sustainability in single-species population models. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1453), 147–162. doi: 10.1098/rstb.2004.1577

Sæther, B.-E., Engen, S., & Lande, R. (1996). Density-Dependence and Optimal Harvesting of Fluctuating Populations. *Oikos*, *76*(1), 40–46. JSTOR. doi: 10.2307/3545746

Sæther, B.-E., Engen, S., & Solberg, E. J. (2001). Optimal harvest of age-structured populations of moose Alces alces in a fluctuating environment. *Wildlife Biology*, *7*(3), 171–179. doi: 10.2981/wlb.2001.021

Stevens, B. S., Bence, J. R., Porter, W. F., & Parent, C. J. (2017). Structural uncertainty limits generality of fall harvest strategies for wild turkeys. *The Journal of Wildlife Management*, *81*(4), 617–628. doi: 10.1002/jwmg.21228

Williams, C. K. (2013). Accounting for wildlife life-history strategies when modeling stochastic density-dependent populations: A review. *The Journal of Wildlife Management*, *77*(1), 4–11. doi: 10.1002/jwmg.429

# Supporting information S2: Additional results

Accompanying manuscript: Sustainability of wildlife harvest in stochastic social-ecological systems.

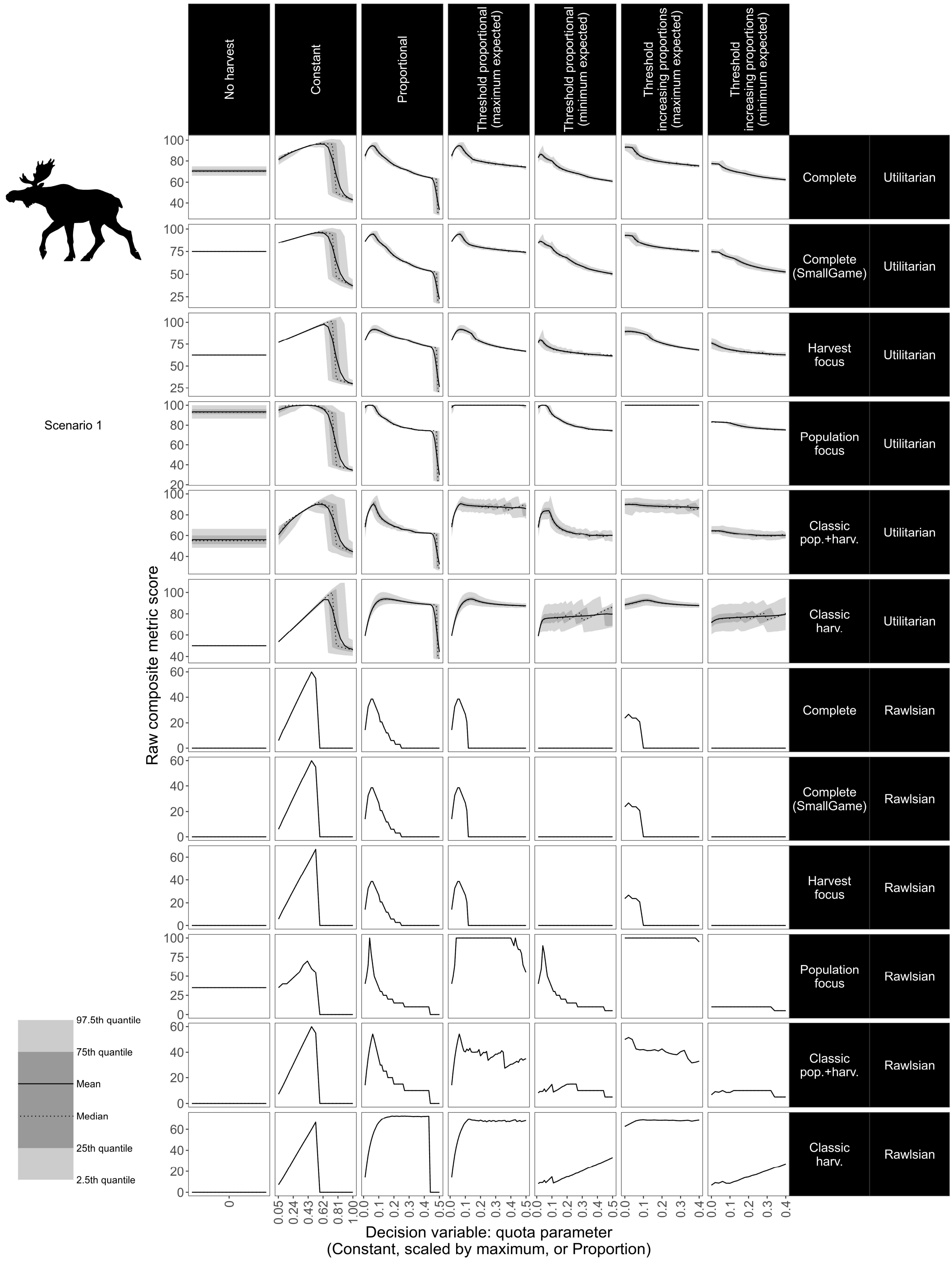
Authors: Elizabeth Law, John D. C. Linnell, Bram van Moorter, Erlend B. Nilsen.

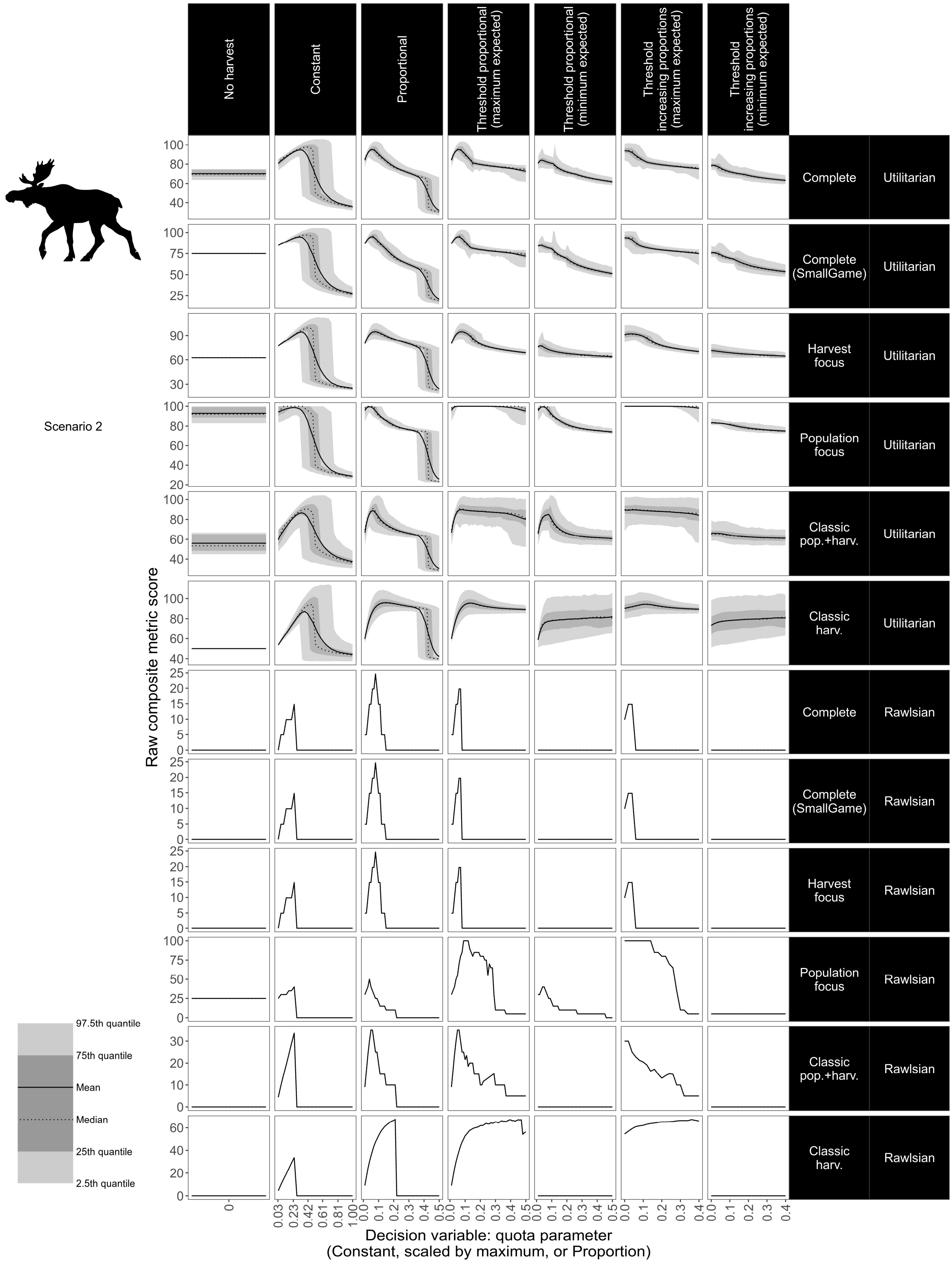
## S2.1 Scores for composite sustainability metrics across decision variables

This supporting information provides variants of the Figure 2 results – composite scores across quota parameter decision variables – for all scenarios, and includes measures of distribution (variability) across iterations.

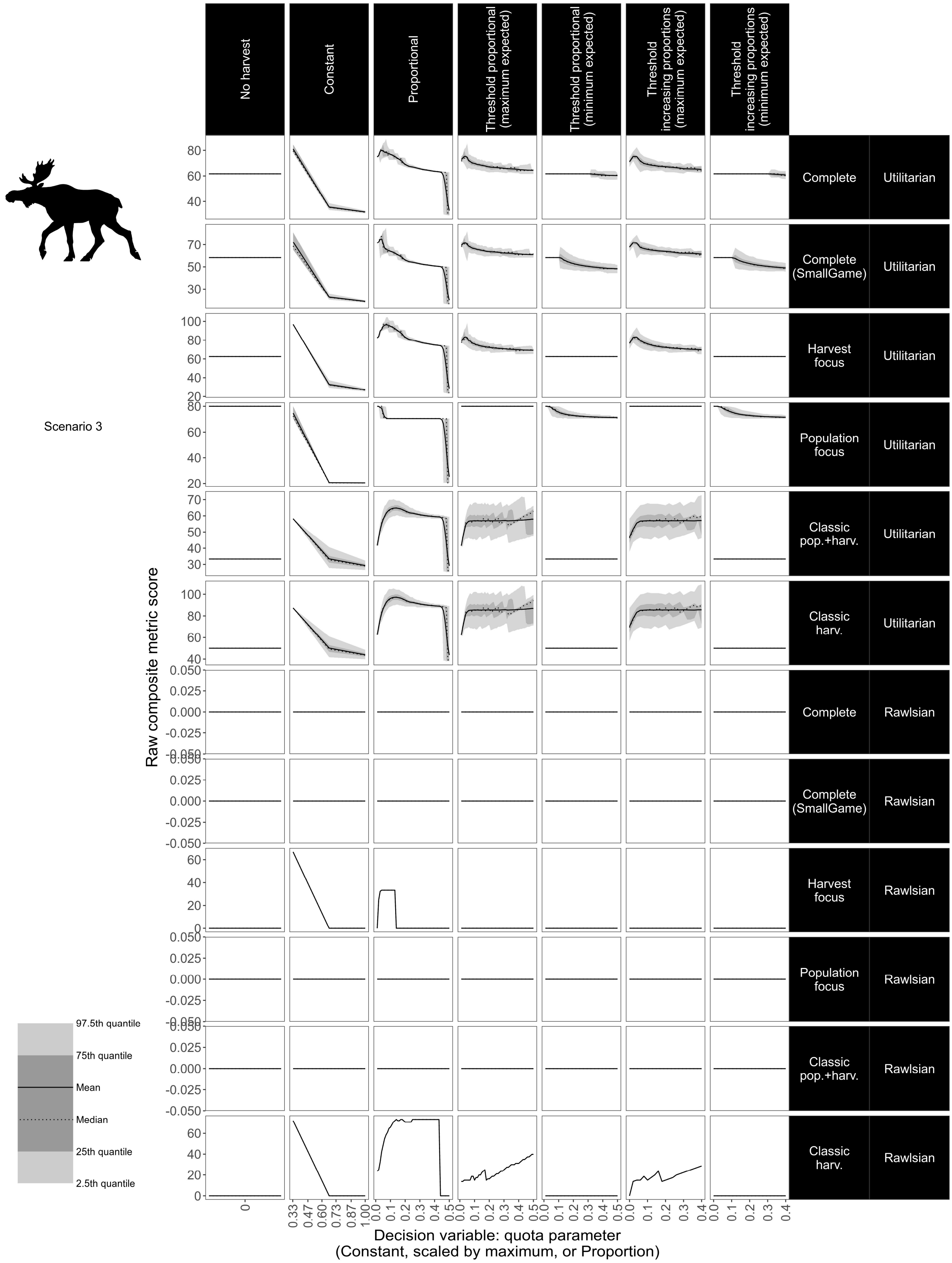
Composite metrics are shown across the quota parameter options available in each harvest strategy. For the threshold-proportional and threshold-increasing proportions, the maximum and minimum expected values are shown (i.e. the maximum and minimum mean values from all the alternative thresholds and gaps between high and low proportions). These are shown individually for each species and start population/variability scenario, including the mean, median, and quantile ranges. Because scores are scaled based on mean scores, quantile ranges above mean scores can be above 100.

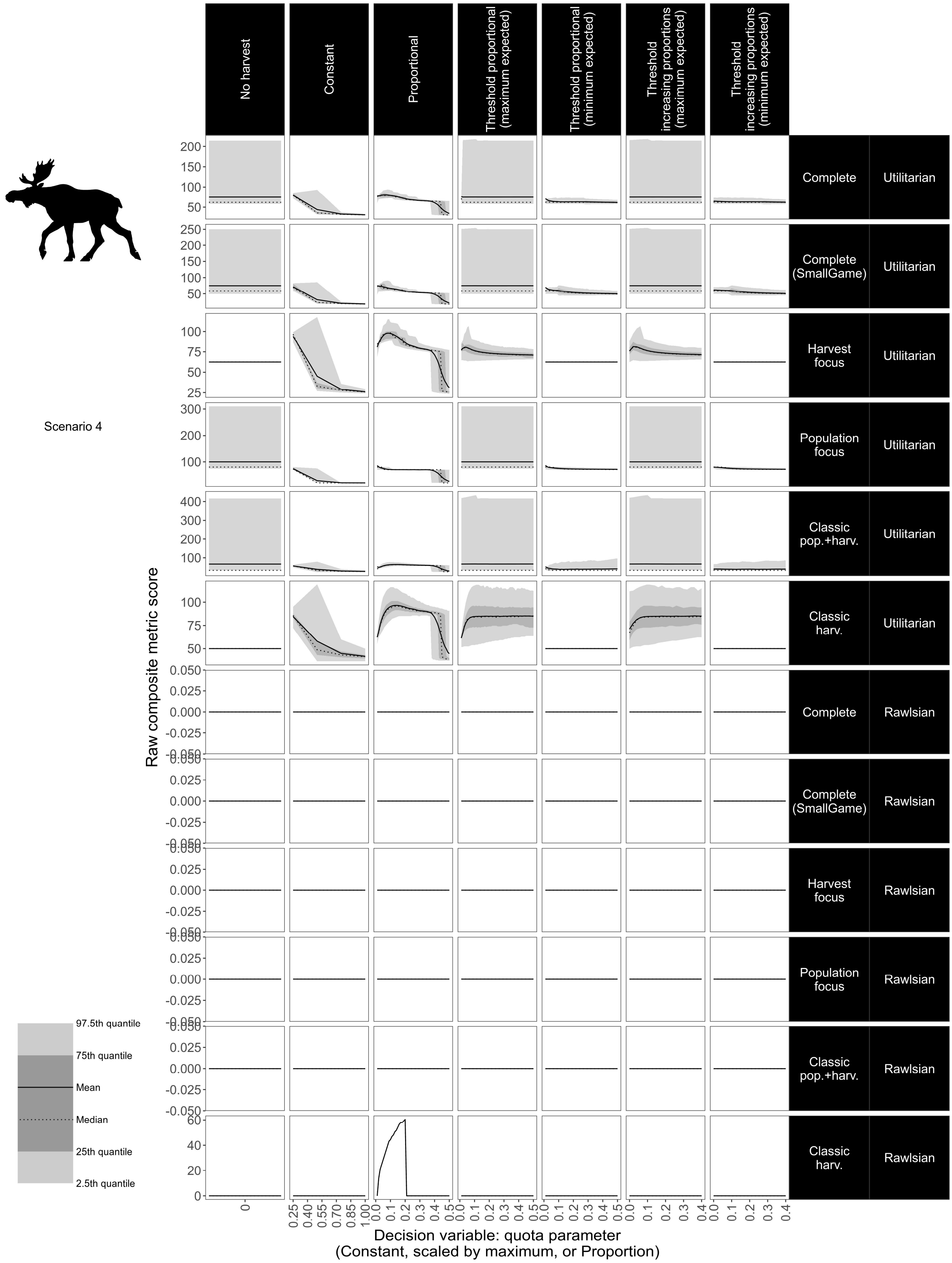
## Figure S2.1.1 Sustainability metrics, Great-unicorn

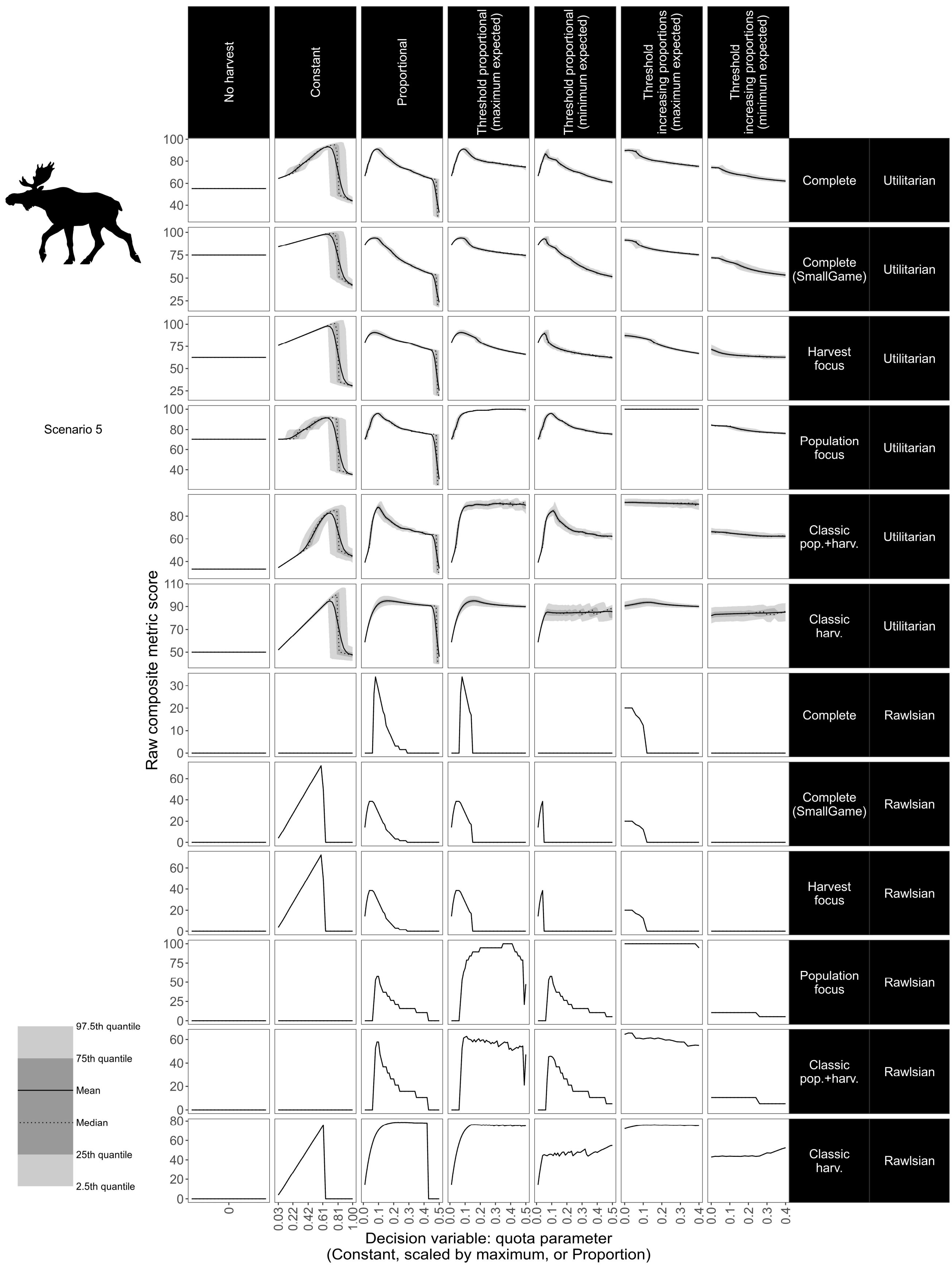


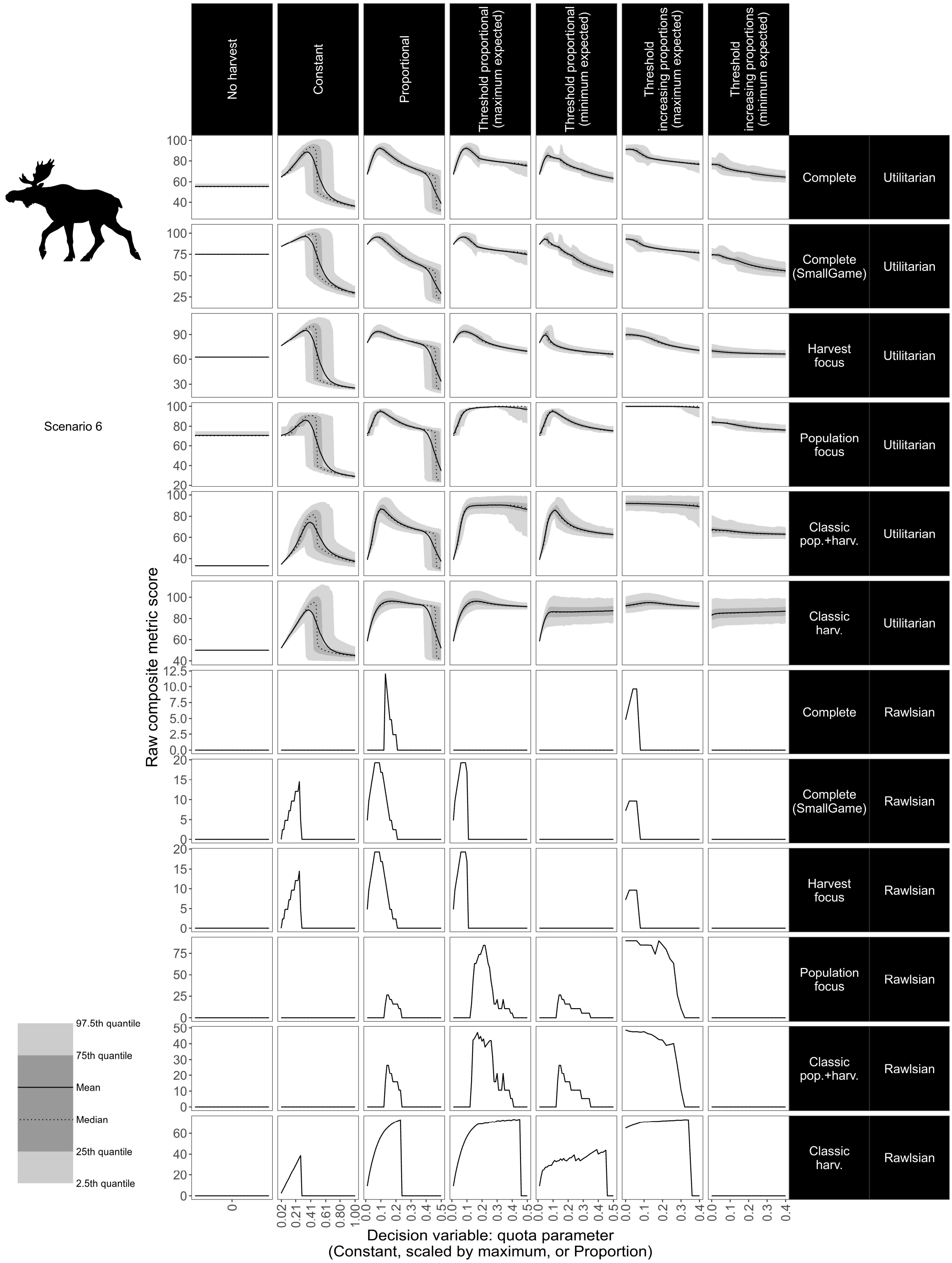


Scenario 2
No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile


No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Scenario 3
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

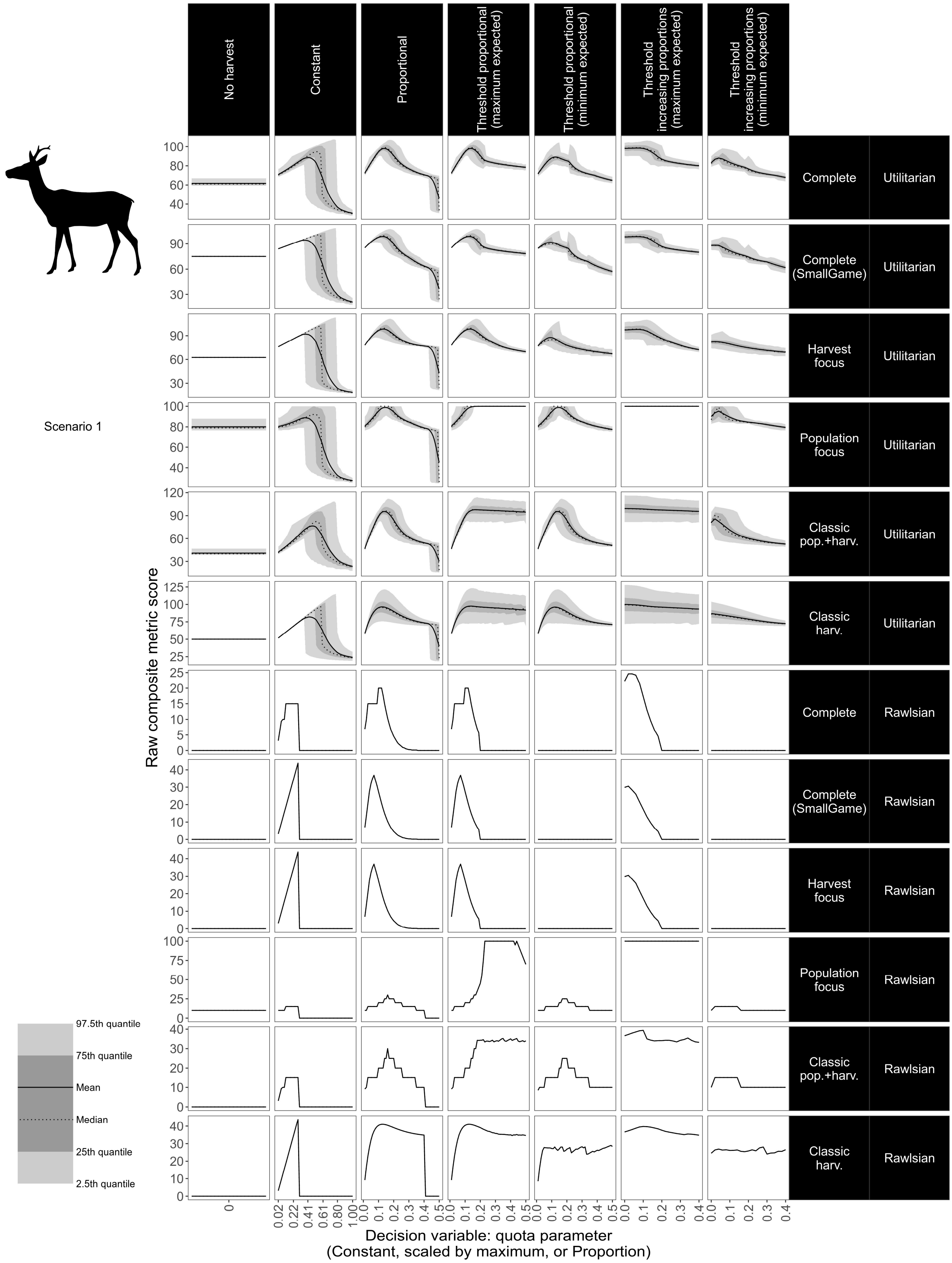

No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Utilitarian
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Rawlsian
Scenario 4
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

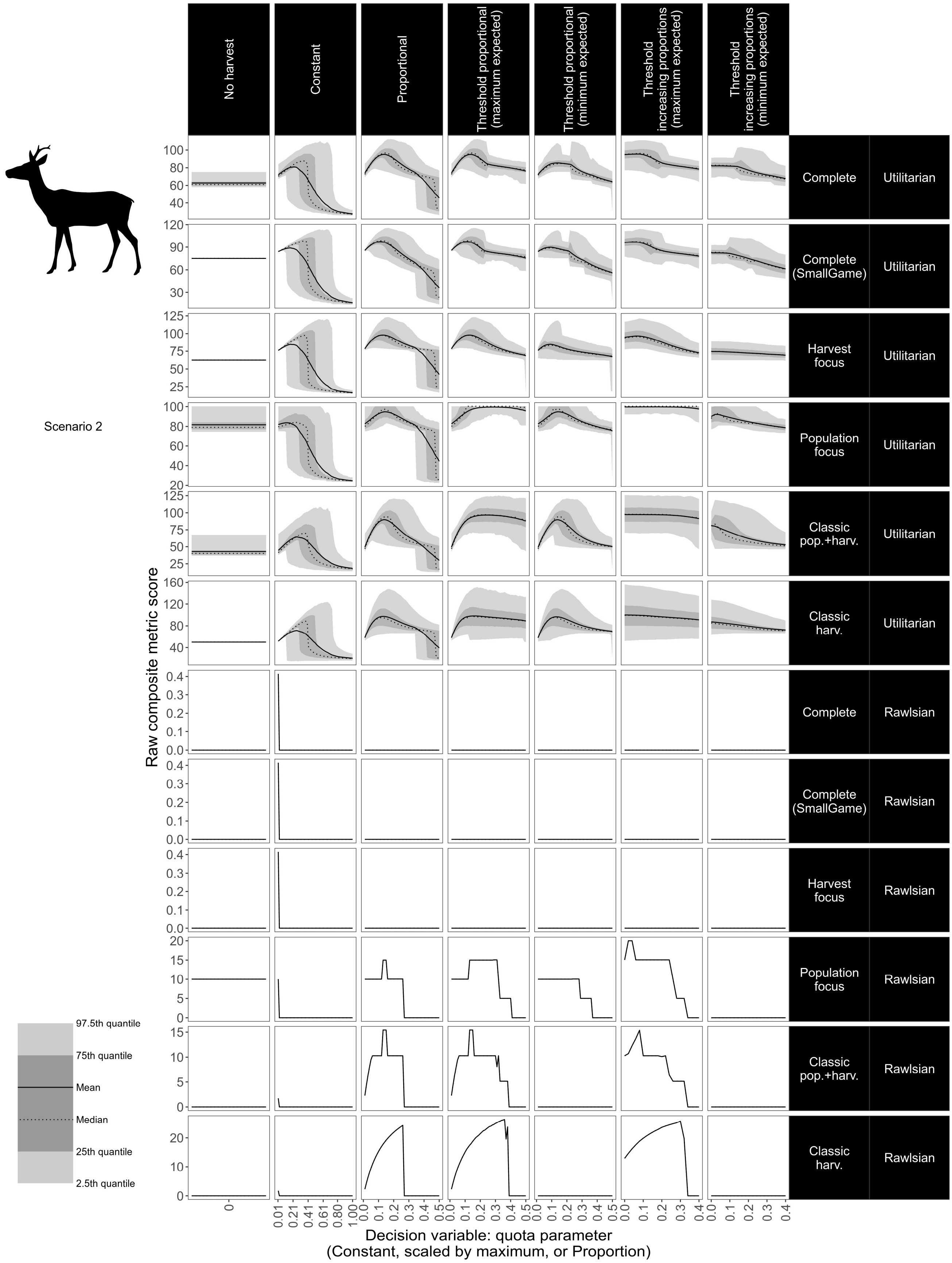

No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Scenario 5
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

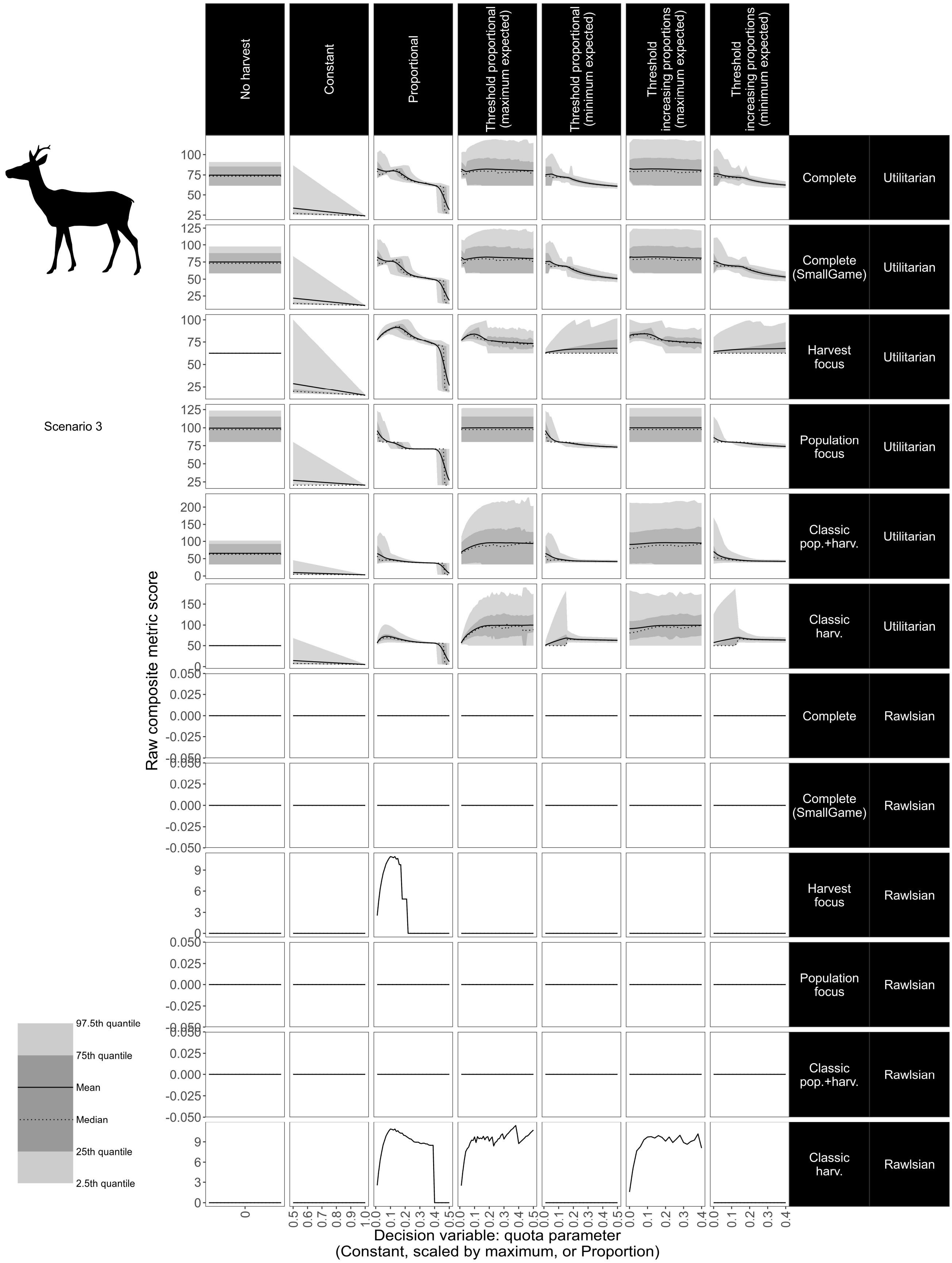


No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Scenario 6
Raw composite metric score
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)

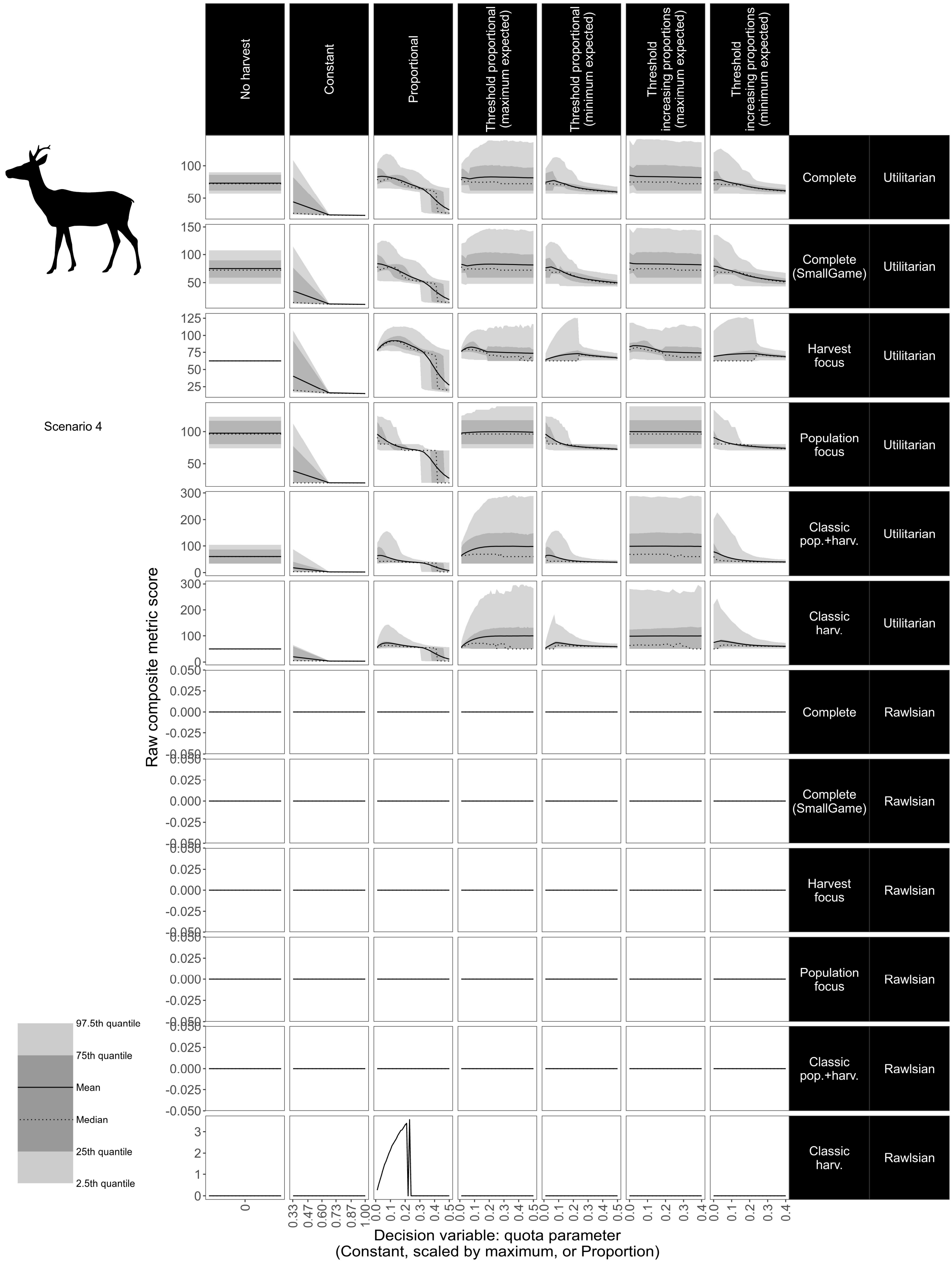
## Figure S2.1.2 Sustainability metrics, Lesser-unicorn



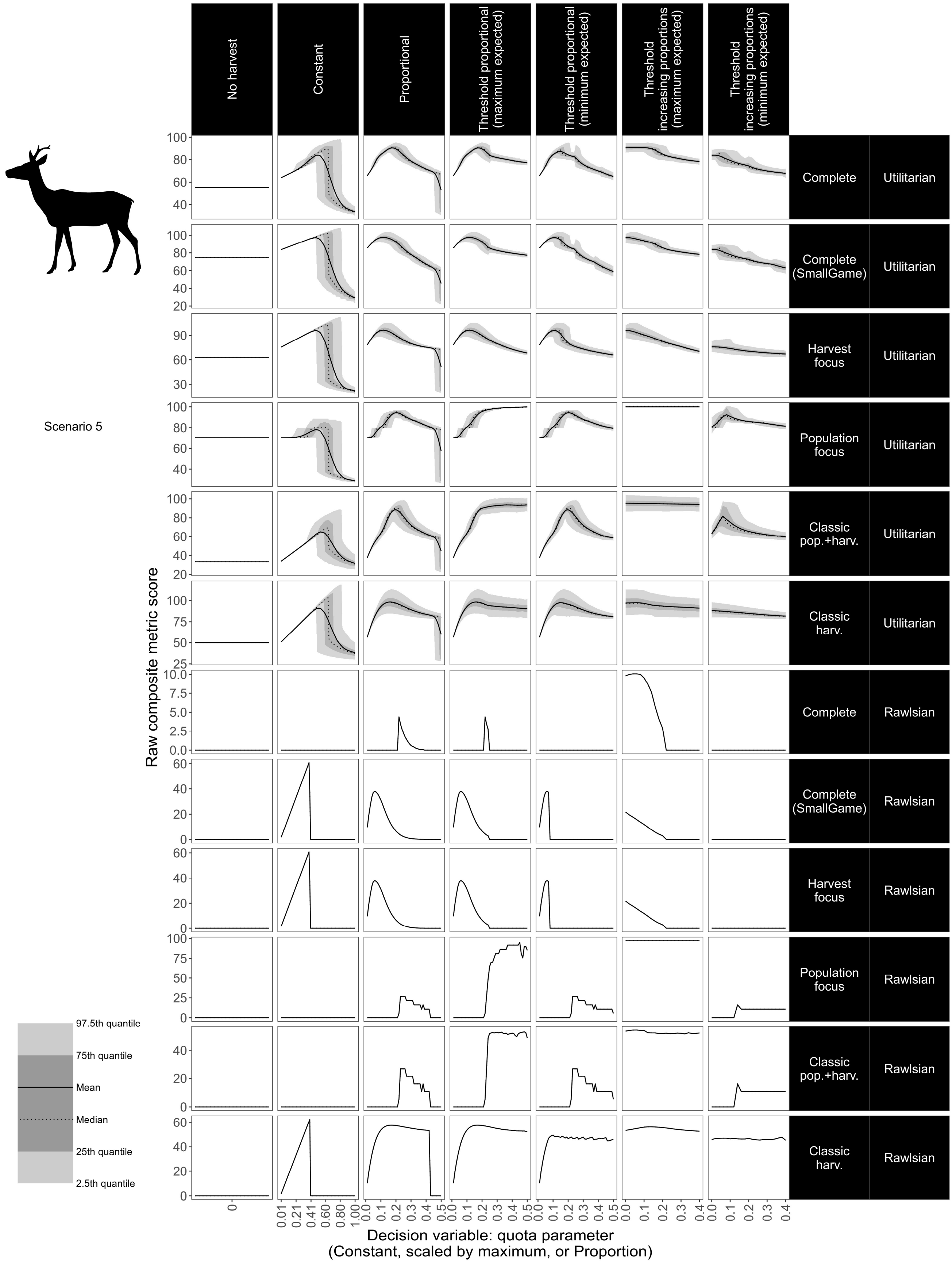

No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Scenario 2
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

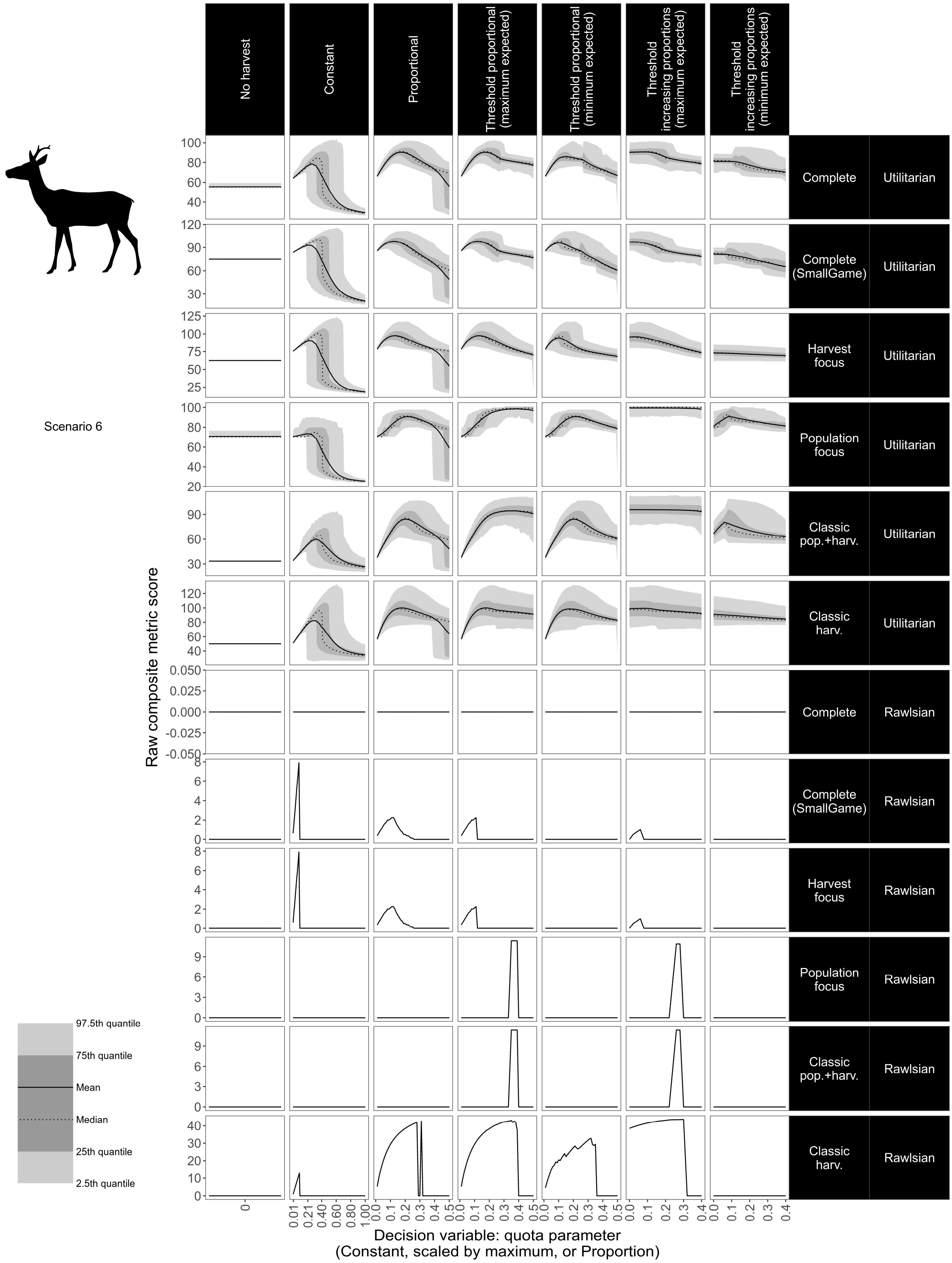


No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Scenario 3
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

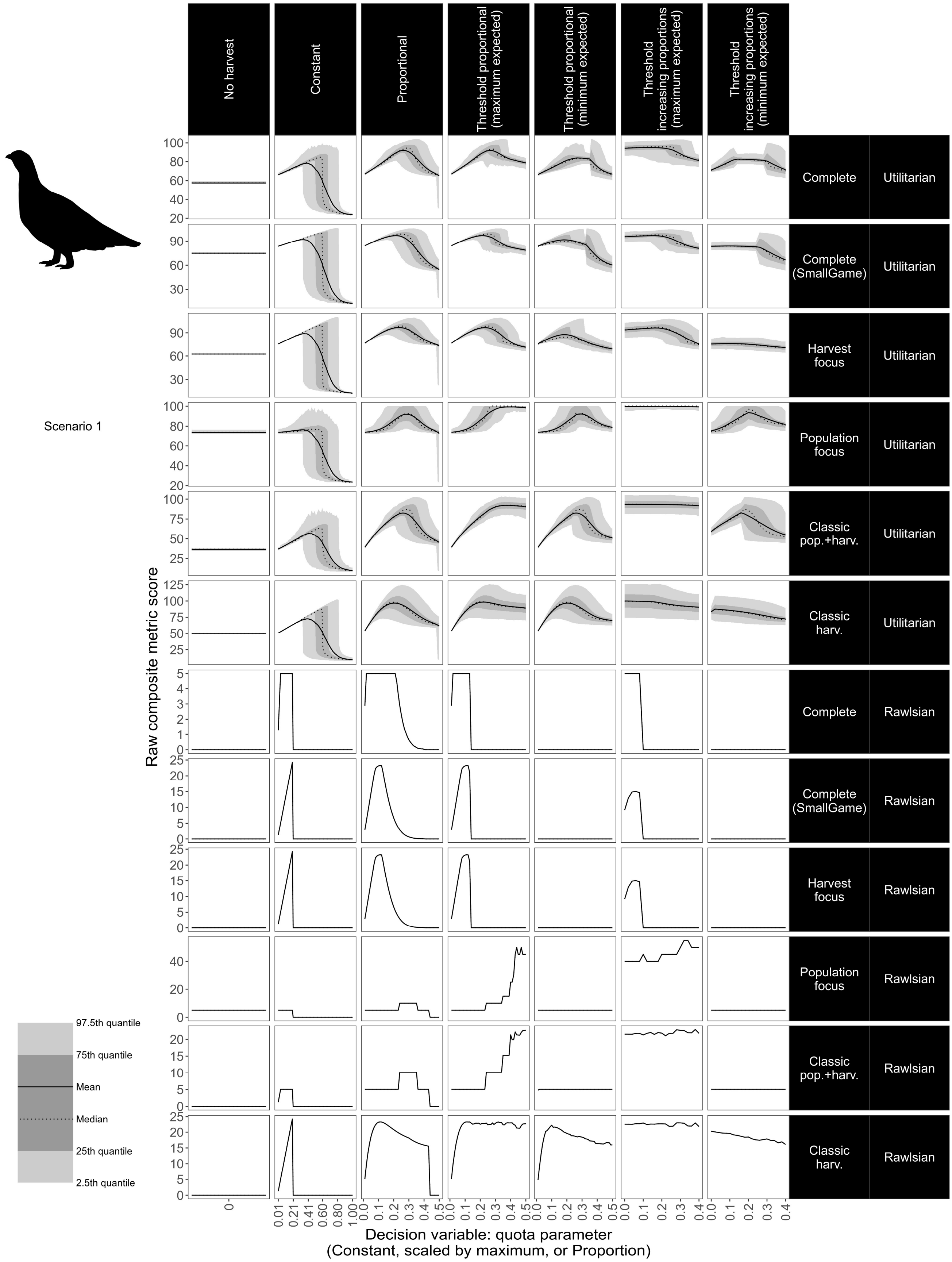


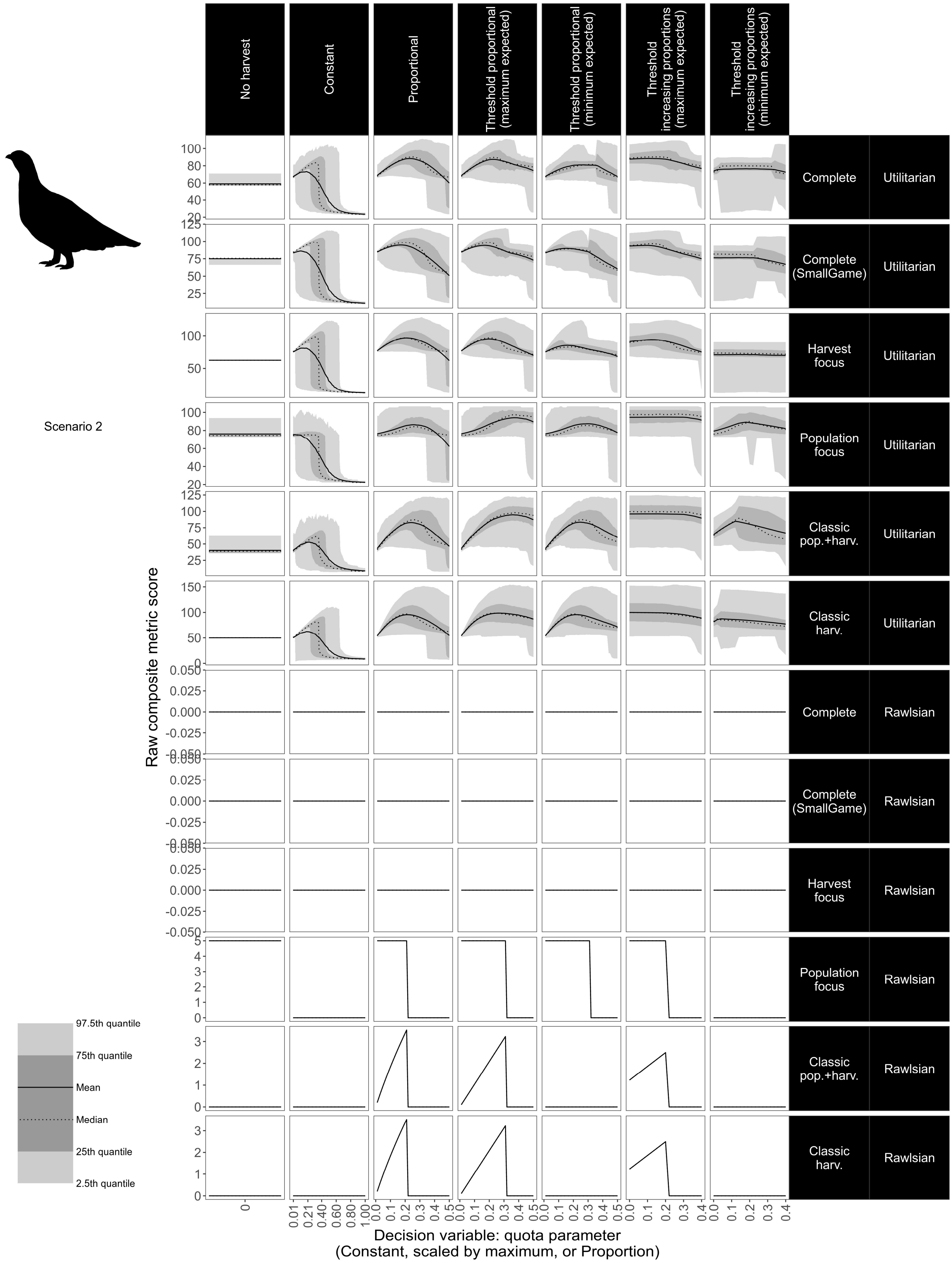
Scenario 4
No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

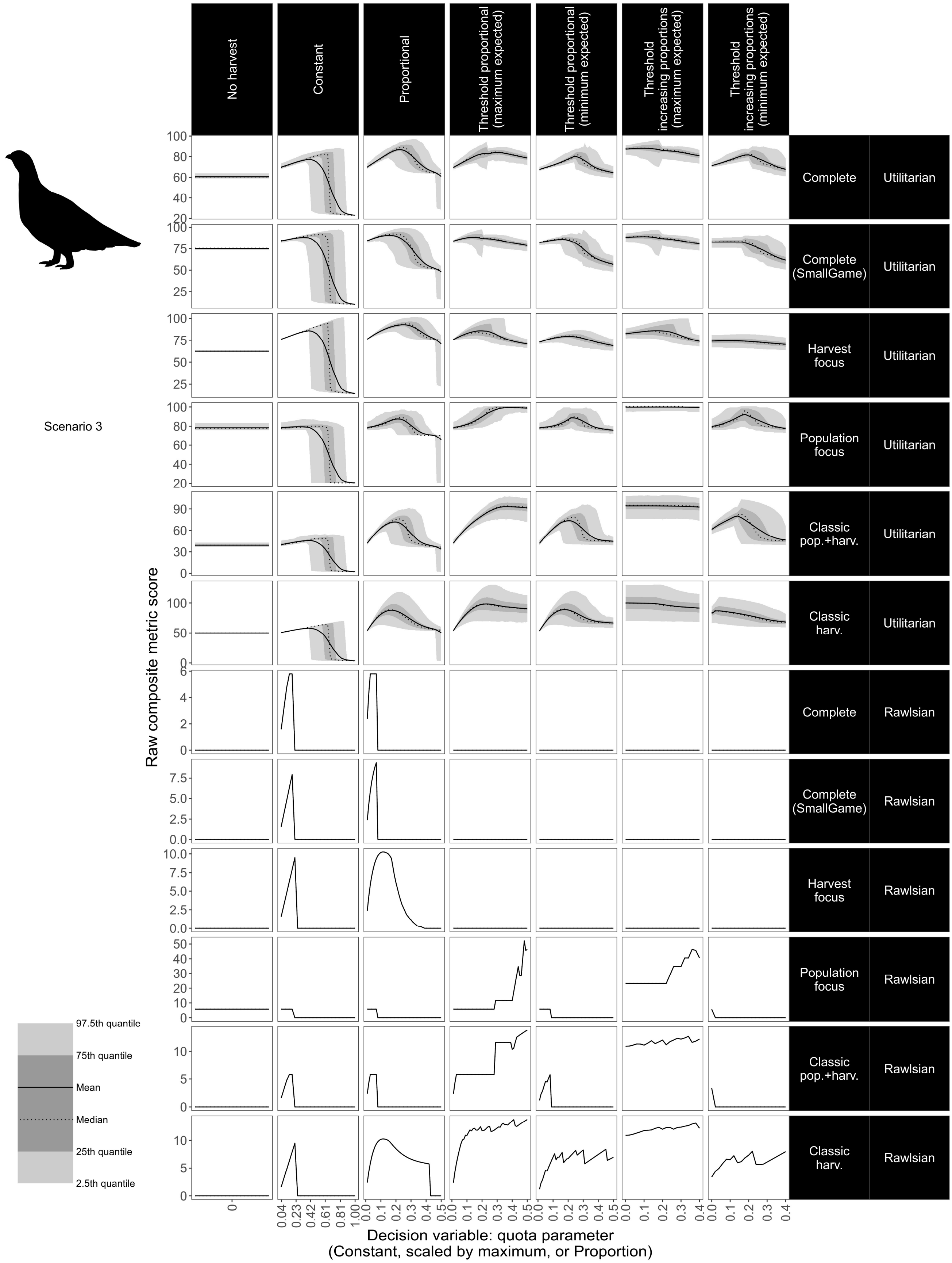

Scenario 5
No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile


No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Utilitarian
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Rawlsian
Scenario 6
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

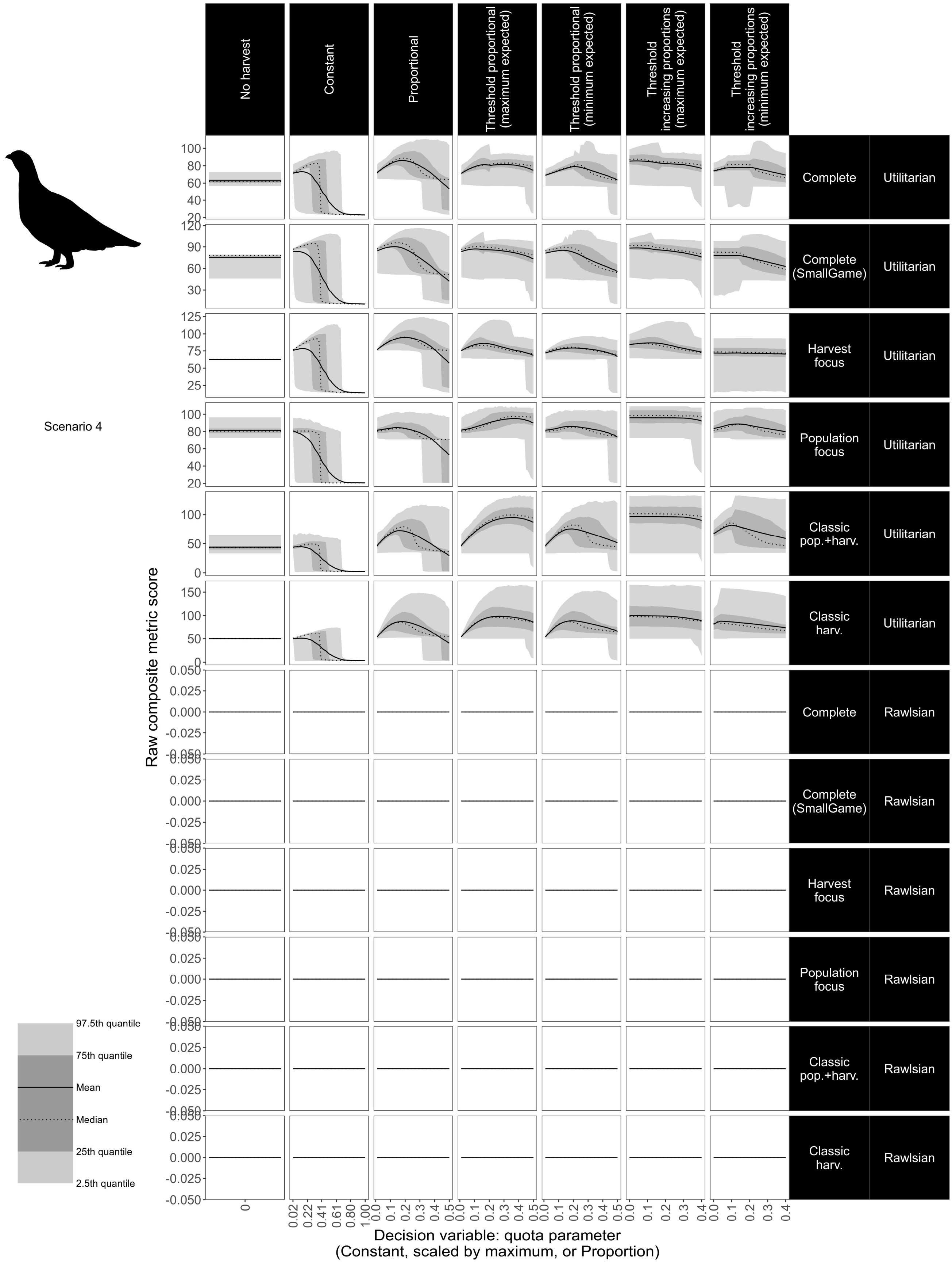
## Figure S2.1.3 Sustainability metrics, Phoenix



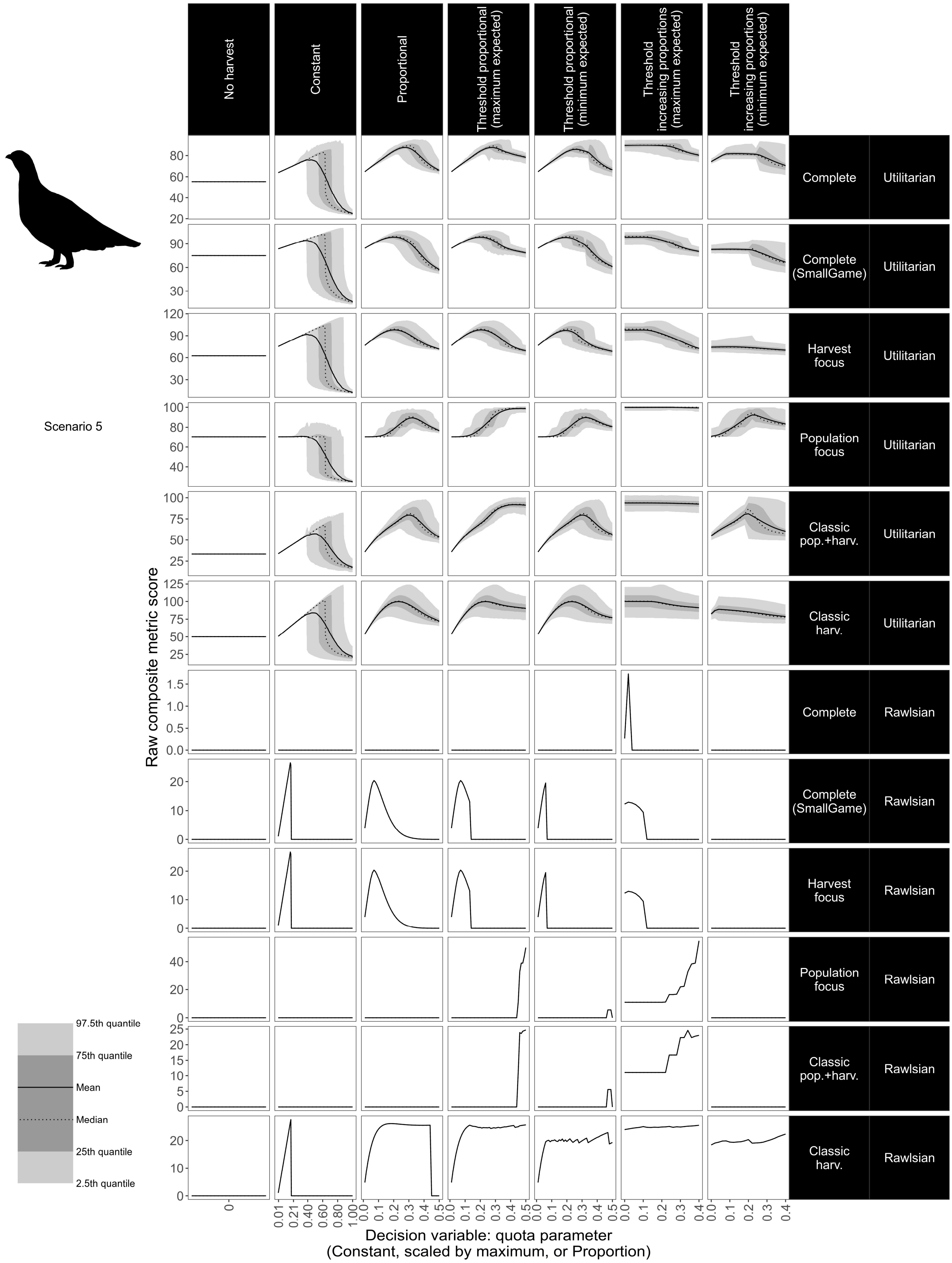

Scenario 2
No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile



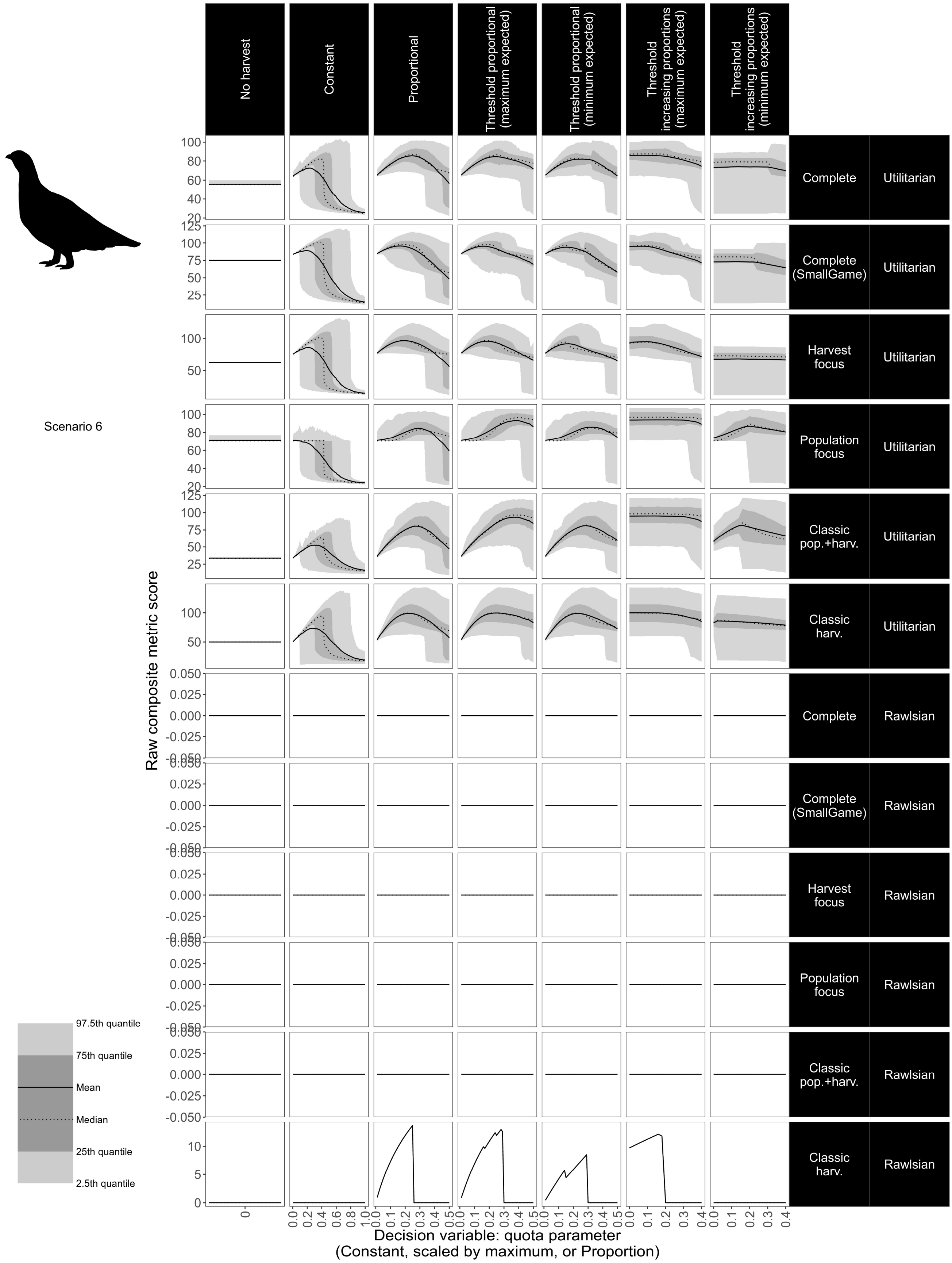
No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Scenario 3
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile



Scenario 4
No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Utilitarian
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Rawlsian
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile


No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Utilitarian
Complete (SmallGame)
Utilitarian
Harvest focus
Utilitarian
Population focus
Utilitarian
Classic pop.+harv.
Utilitarian
Classic harv.
Utilitarian
Complete
Rawlsian
Complete (SmallGame)
Rawlsian
Harvest focus
Rawlsian
Population focus
Rawlsian
Classic pop.+harv.
Rawlsian
Classic harv.
Rawlsian
Scenario 5
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

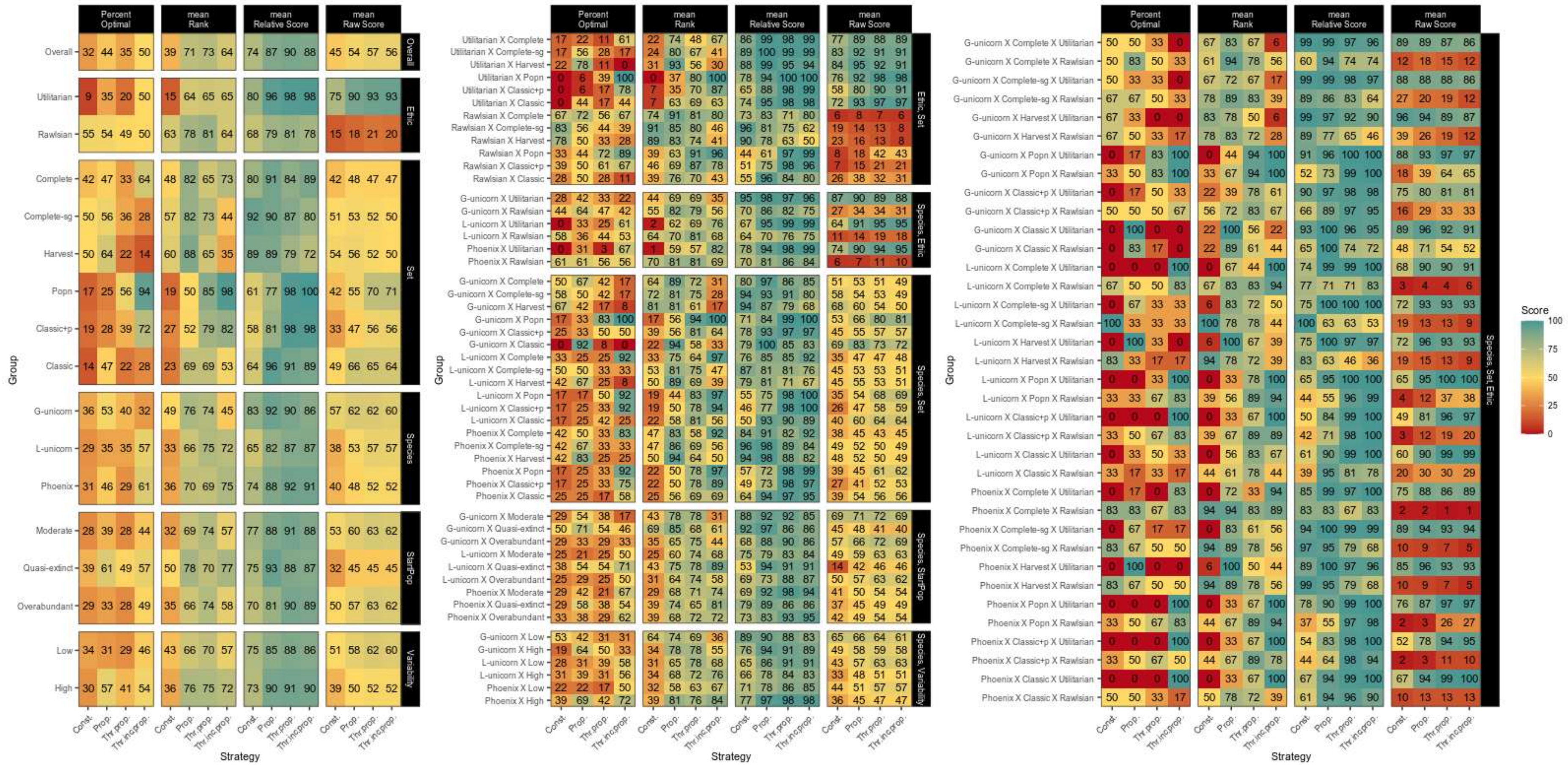
1054

1055 
No harvest
Constant
Proportional
Threshold proportional (maximum expected)
Threshold proportional (minimum expected)
Threshold increasing proportions (maximum expected)
Threshold increasing proportions (minimum expected)
Complete
Complete (SmallGame)
Harvest focus
Population focus
Classic pop.+harv.
Classic harv.
Utilitarian
Rawlsian
Scenario 6
Raw composite metric score
Decision variable: quota parameter
(Constant, scaled by maximum, or Proportion)
97.5th quantile
75th quantile
Mean
Median
25th quantile
2.5th quantile

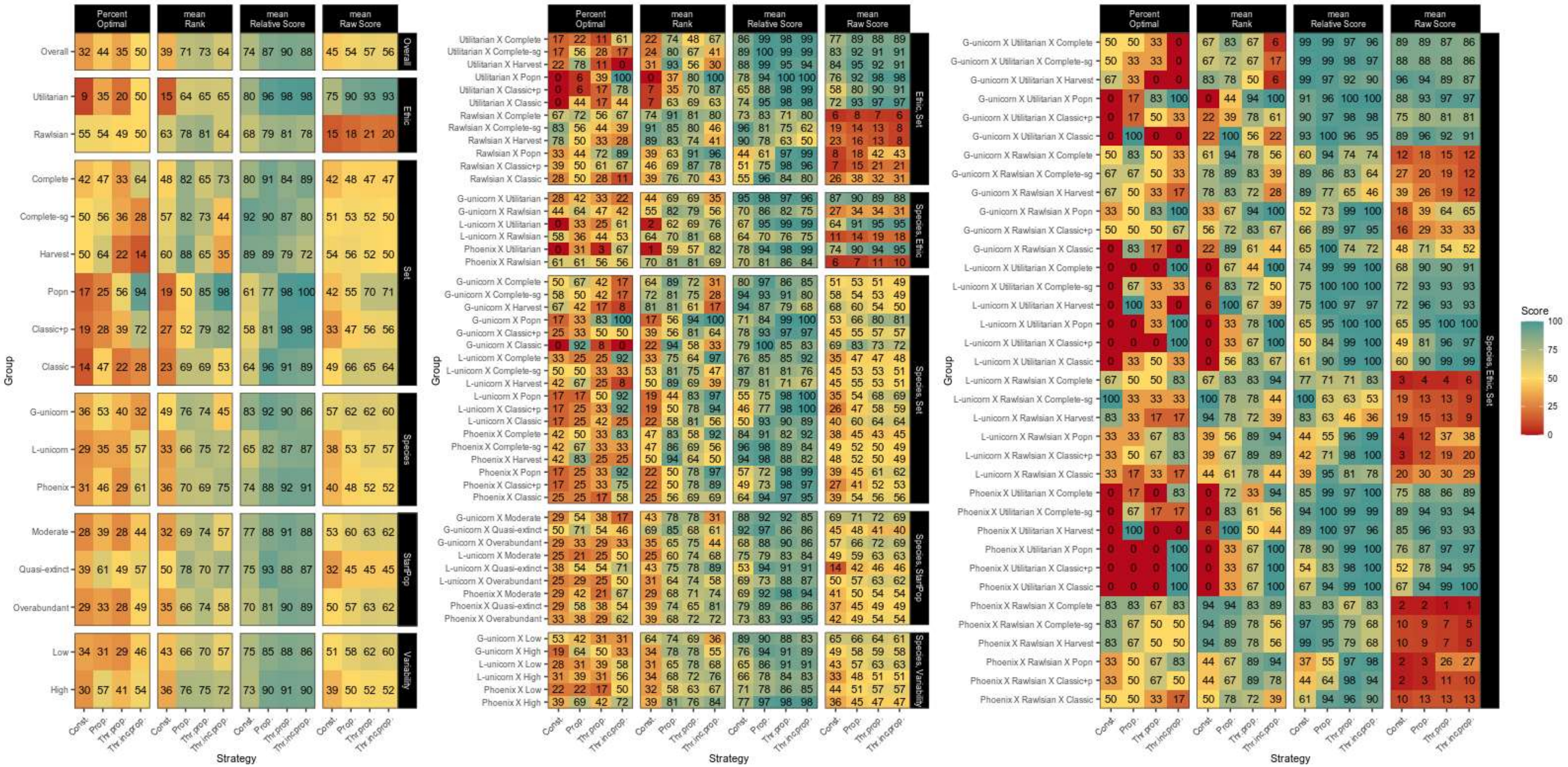
## S2.2 Percentage optimal, and mean rank and scores, by strategy

### Figure S2.2.1 Percentage optimal, rank, and mean scores

These figures show, within the different context groups, the percentage of cases a strategy is considered optimal (i.e. having the highest score; note these will not add to 100 as multiple strategies can have the same score and thus be jointly optimal), the mean rank (i.e. ranked by score, rescaled to 0:100, with 100 being best), the mean relative score (out of a maximum of 100, showing relative performance against other strategies), and mean raw score (showing perceived performance of the strategy). The first set of panels show overall and single factor groups, the second set selected two-factor groups, and the third set select three-factor groups. Selected groups focus on Species, Ethic and Set. See below for a breakdown of multi-strategy optimality.

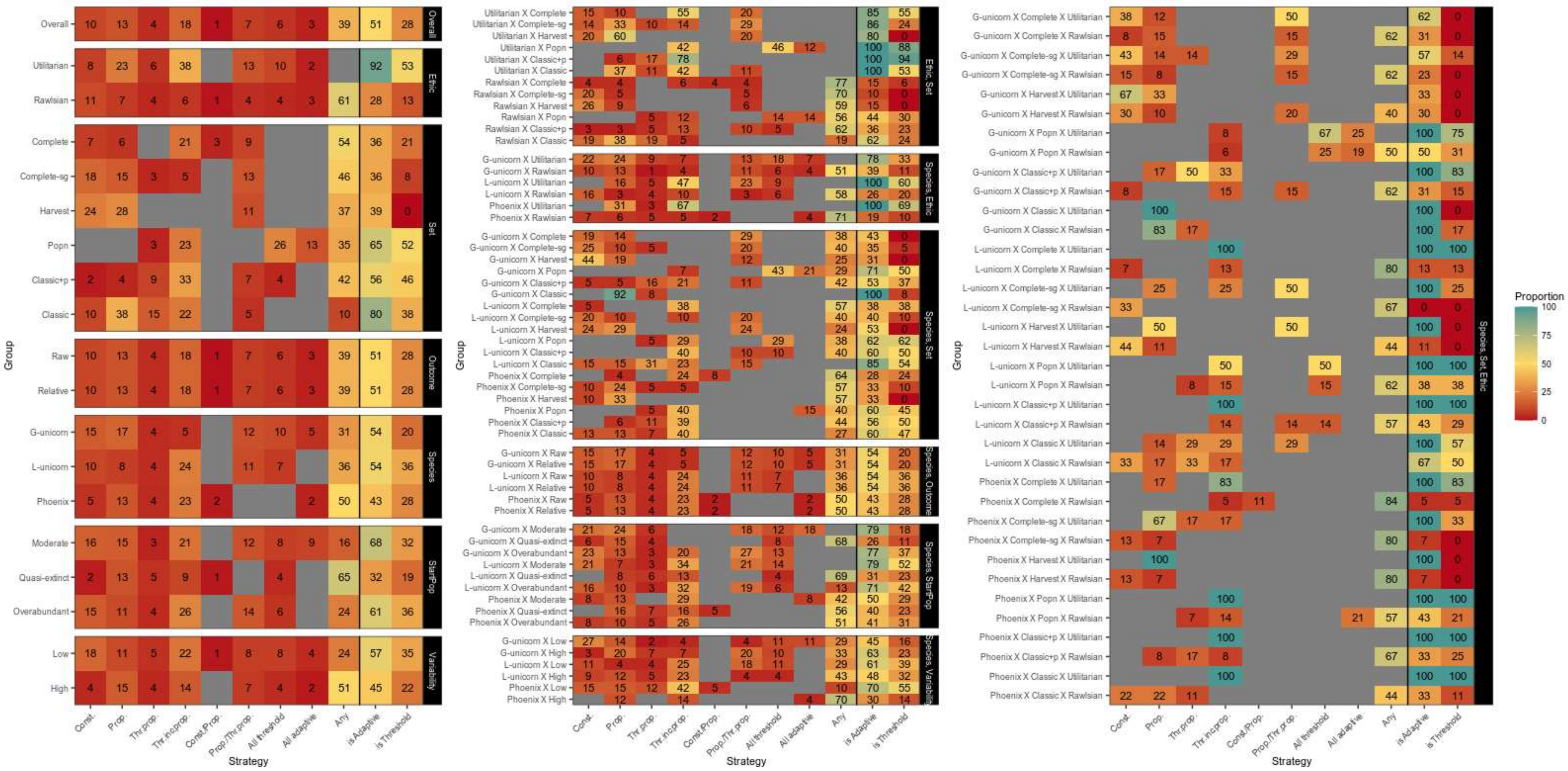


This is the same as above, with different ordering of the three panel factors.

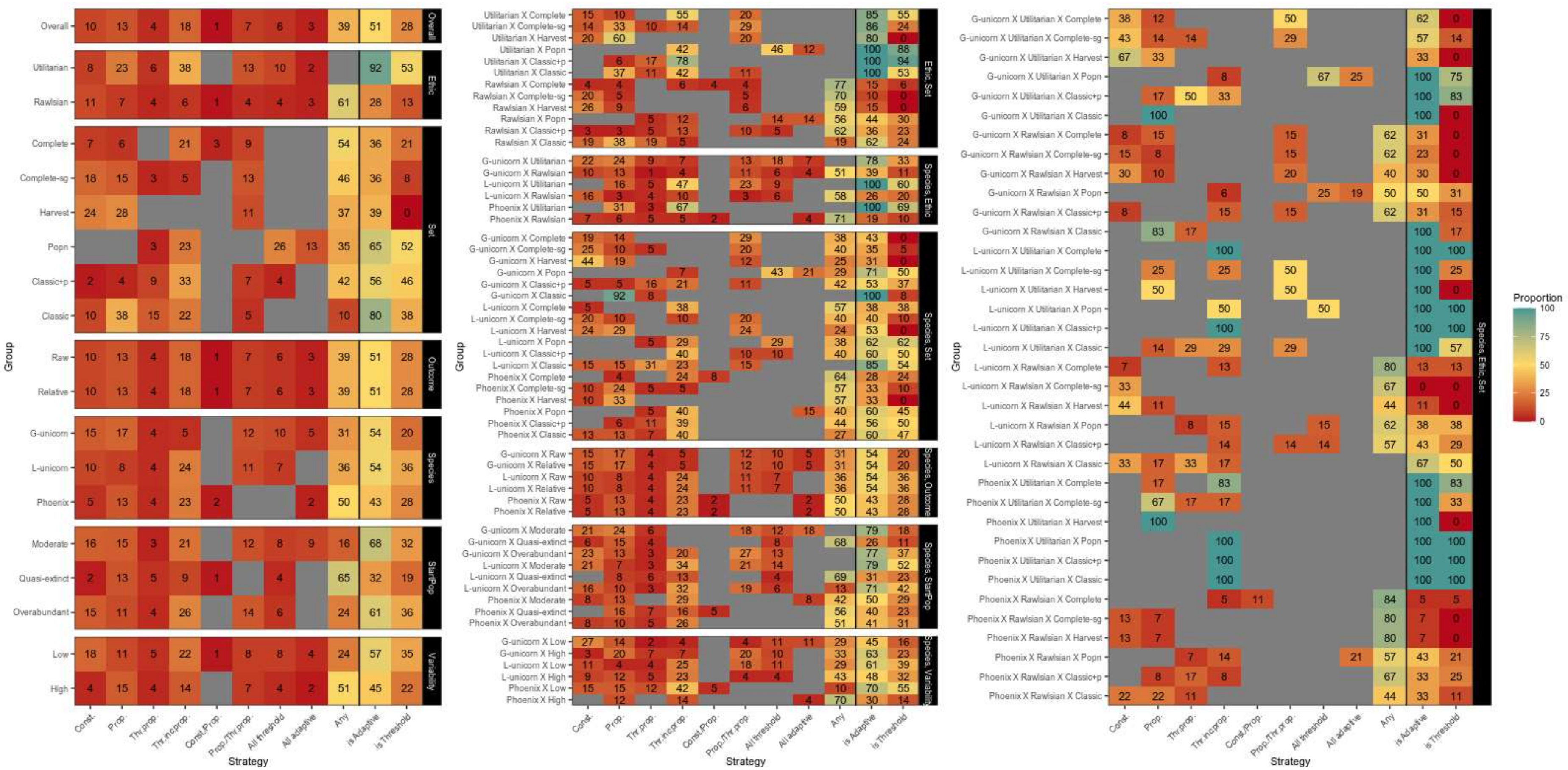


## Figure S2.2.2 Percentage optimal strategy or strategies

These figures show, within the different context groups, the percentage of cases a strategy (or multiple strategies) is considered optimal. Two summary columns are included: ‘is Adaptive’, the percentage of cases where the optimal strategy is one or more of the adaptive strategies (i.e. proportional, threshold-proportional, or threshold-increasing-proportions), and ‘is Threshold’, the percentage of cases where the optimal strategy is one or more of the threshold-based strategies (i.e. threshold-proportional or threshold-increasing-proportions). The first set of panels show overall and single factor groups, the second set selected two-factor groups, and the third set select three-factor groups. Here, instances where a strategy group is not ever optimal are allocated NA (grey) whereas other scores are rounded to the nearest integer.

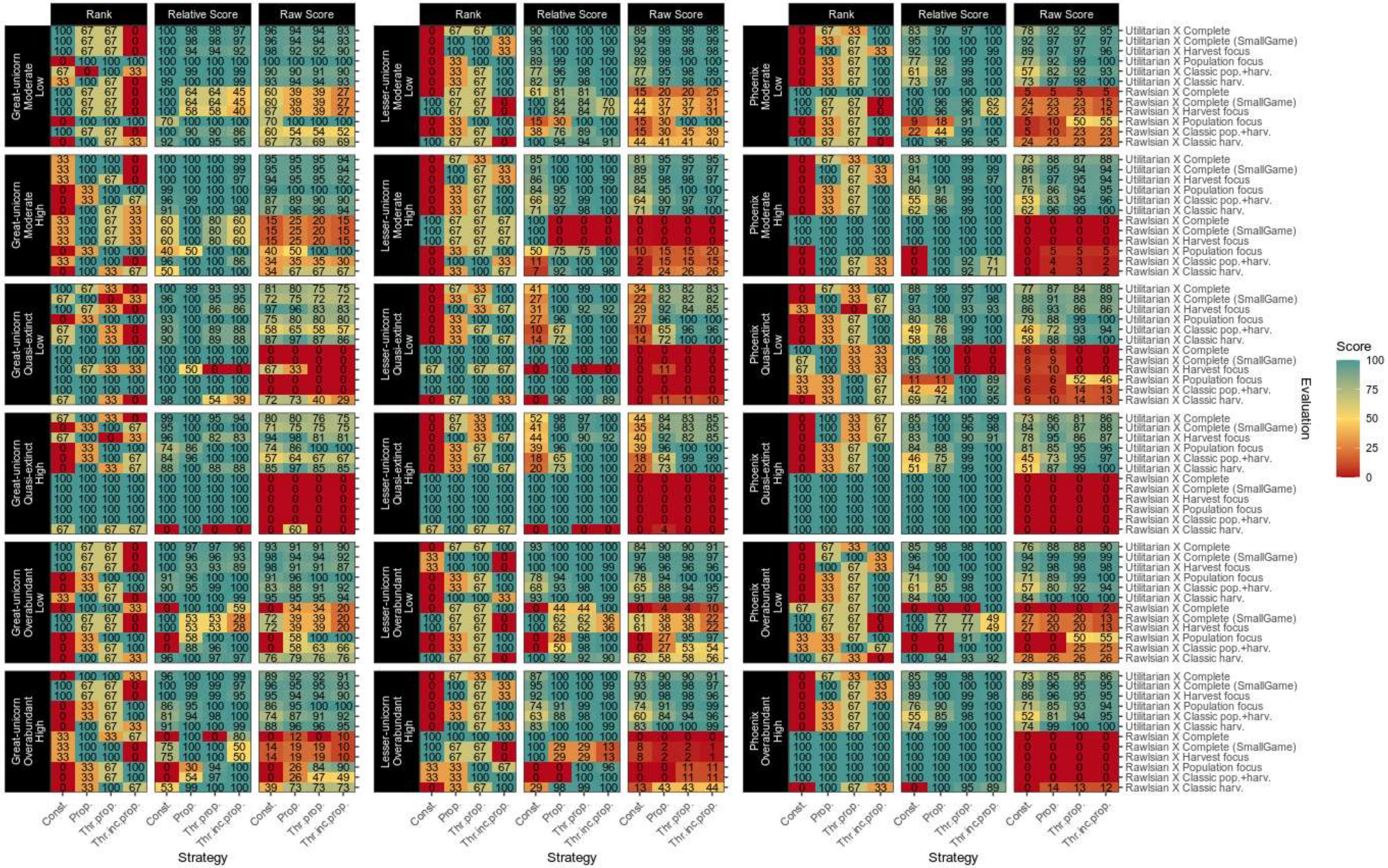


And again with different ordering of the three-factor categories:



## Figure S2.2.3 Environmental contexts by evaluation context

Individual scores with rank (panel set columns), for each harvest strategy (panel columns), for each environmental context (species scenario, panel set rows) and evaluation contexts (panel rows). This shows how, in every environmental context, every harvest strategy could be viewed as optimal, or very close to. However, for constant harvests, optimality is only found in the Rawlsian ethics for the faster life history species (and very low starting populations with high variability in the Great-unicorn) – and often a result of an uninformative metric (i.e. all strategies score a zero in this metric).



## S2.3 Conditional inference trees for optimal harvest strategy and perceived sustainability

Conditional inference trees for optimal harvest strategy and perceived sustainability. In optimal harvest trees, where multiple strategies give equally good (or bad) outcomes, these are allocated to multi-strategy classes.

*Codes:*

**Ethic:** Utilitarian = Utilitarian, Rawlsian = Rawlsian

**Set:** Complete = Comp, Complete (SmallGame) = Complete-sg = Comp-s, Harvest focus = Harvest = Harv, Population focus = Popn, Classic pop.+harv. = Classic+p = Clp+h, Classic harv. = Classic = Clh

**Outcome:** Raw = Raw, Relative = Relative

**Species:** Great-unicorn = G-unicorn = GU, Lesser-unicorn = L-unicorn = LU, Phoenix = Phoenix = Phx

**Variability:** High = High, Low = Low

**Starting population (StartPop):** Moderate = Mod, Quasi-extinct = Low, Overabundant = High

**Strategy:** Constant = Const. = Cnst, Proportional = Prop. = Prop, Threshold-proportional = Thr.prop. = TP, Threshold-increasing-proportions = Thr.inc.prop. = TIP

**Perceived outcome:** Good = score of 85 ≤ 100, Bad = score of 0 ≤ 85.

Figure S2.3.1 Optimal harvest strategy, all species

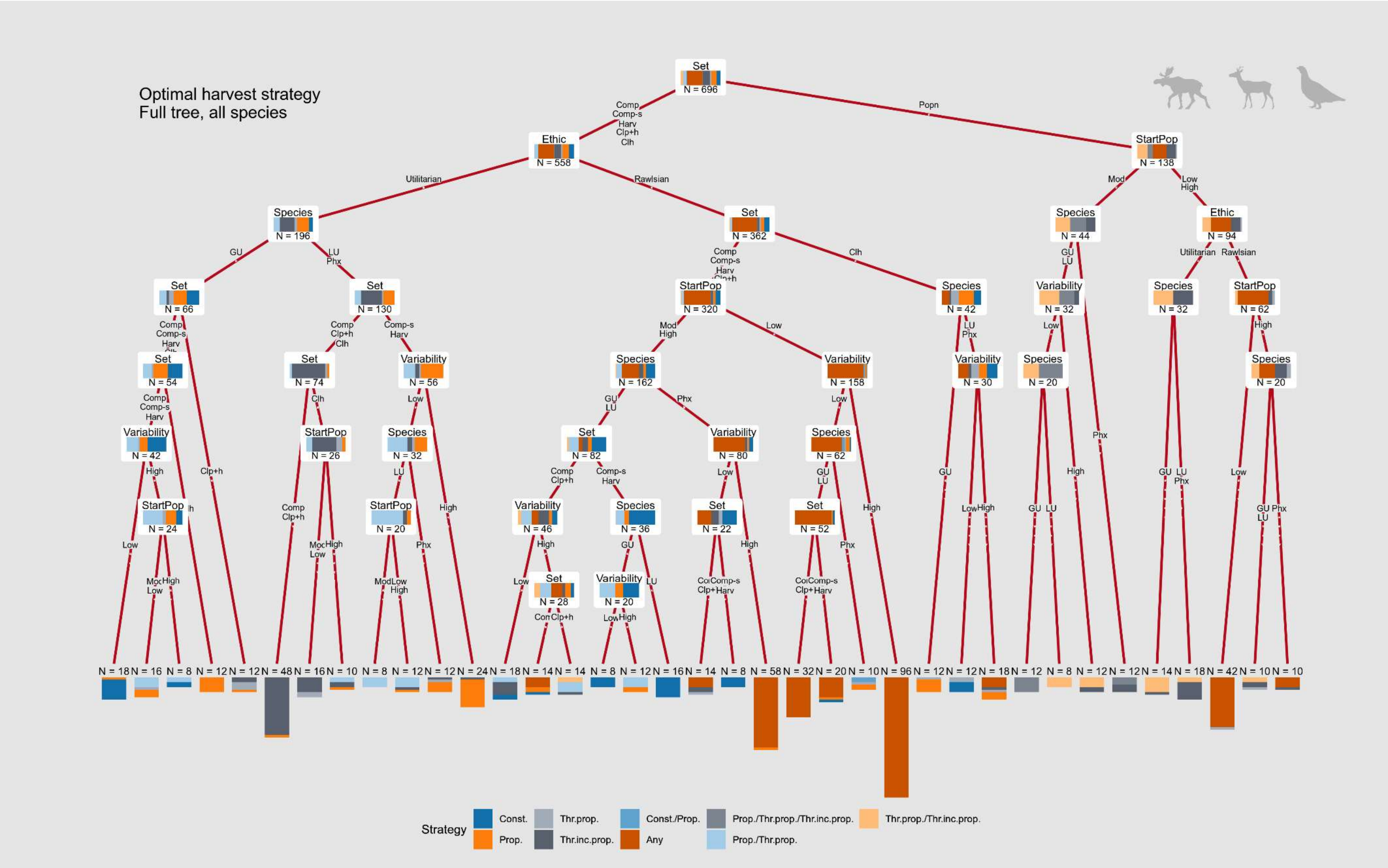
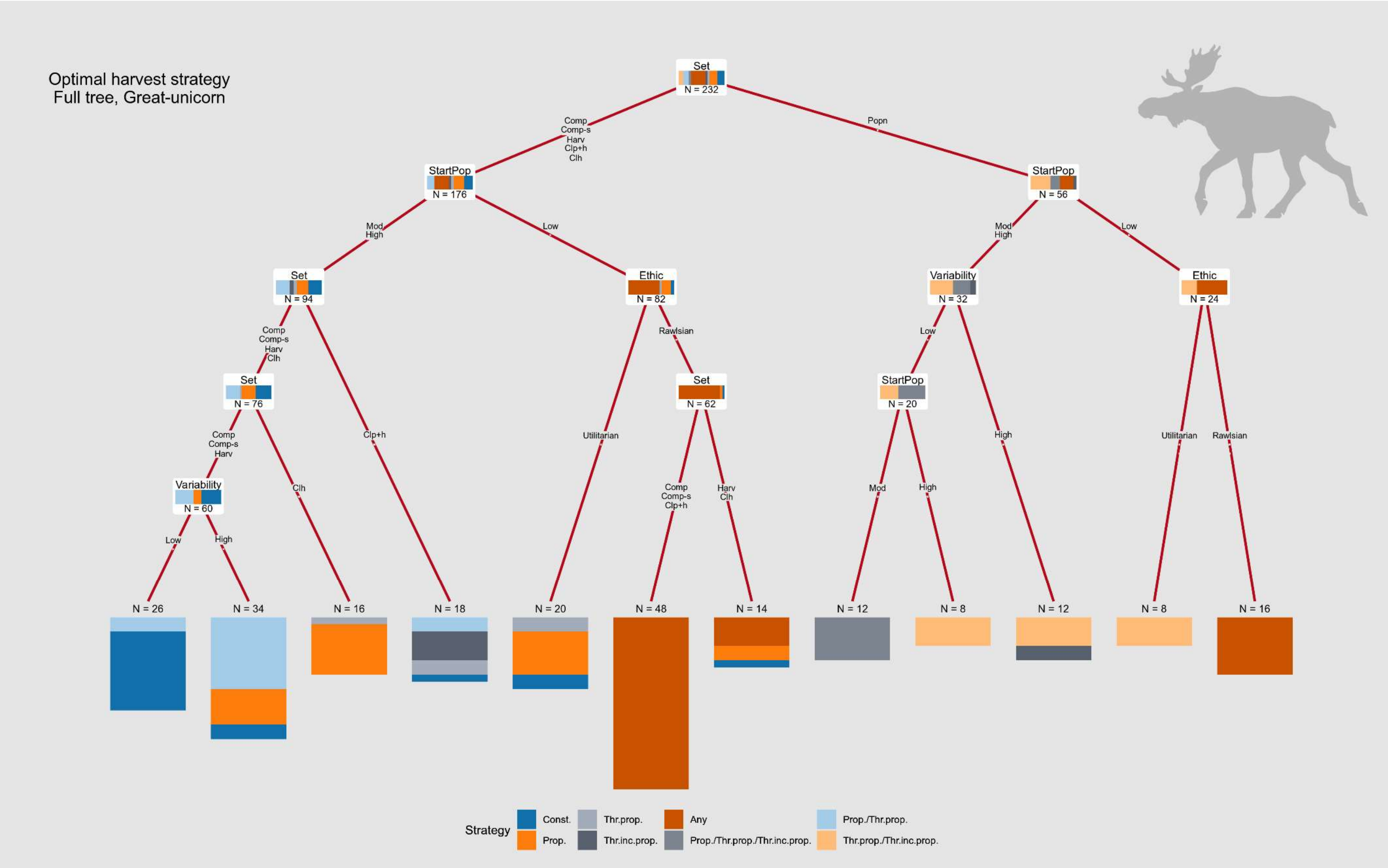
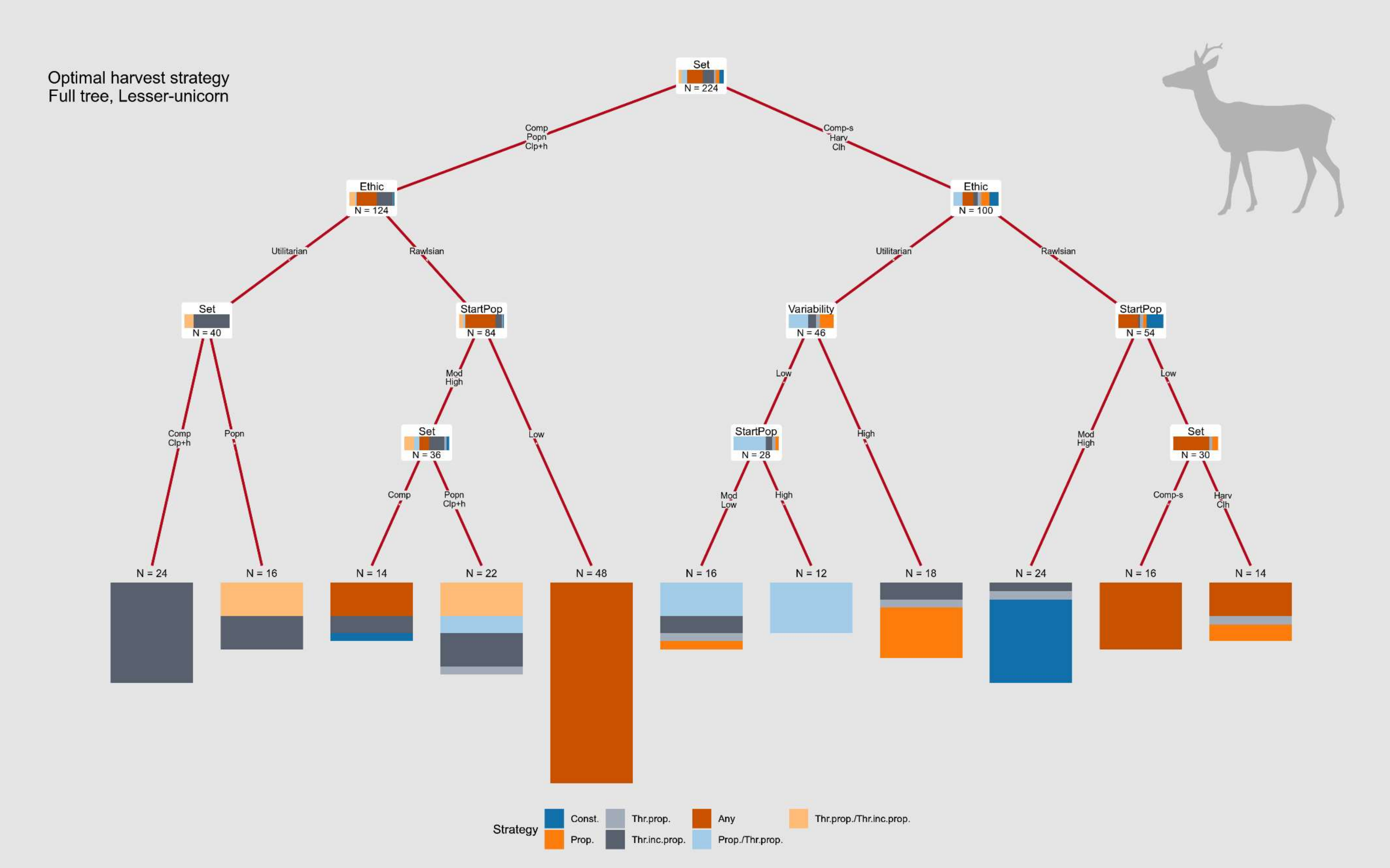


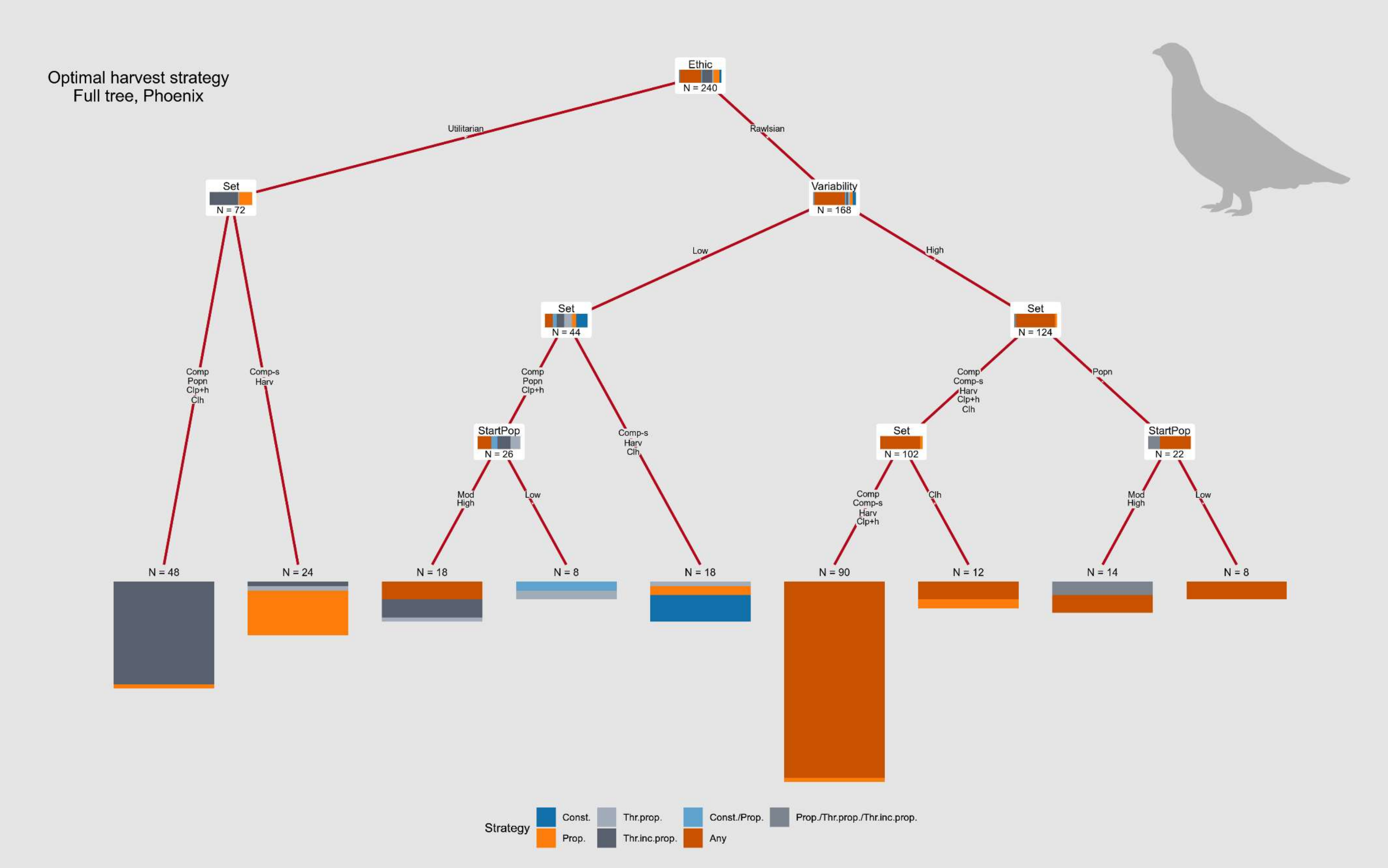
Figure S2.3.2 Optimal harvest strategy, Great-unicorn



## Figure S2.3.3 Optimal harvest strategy, Lesser-unicorn



1108 Figure S2.3.3 Optimal harvest strategy, Phoenix




1109

1110

Figure S2.3.4 Perceived sustainability, all species

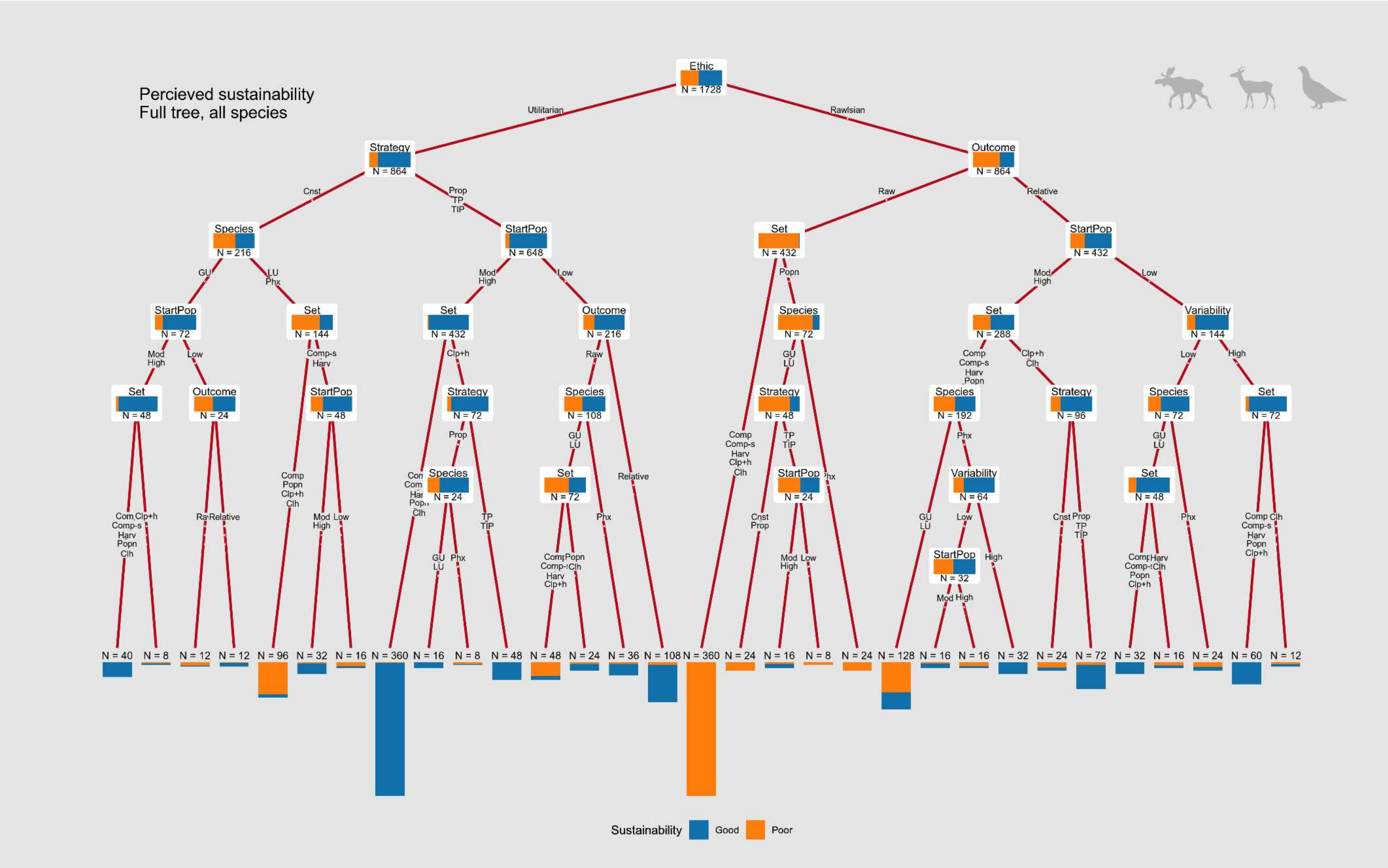
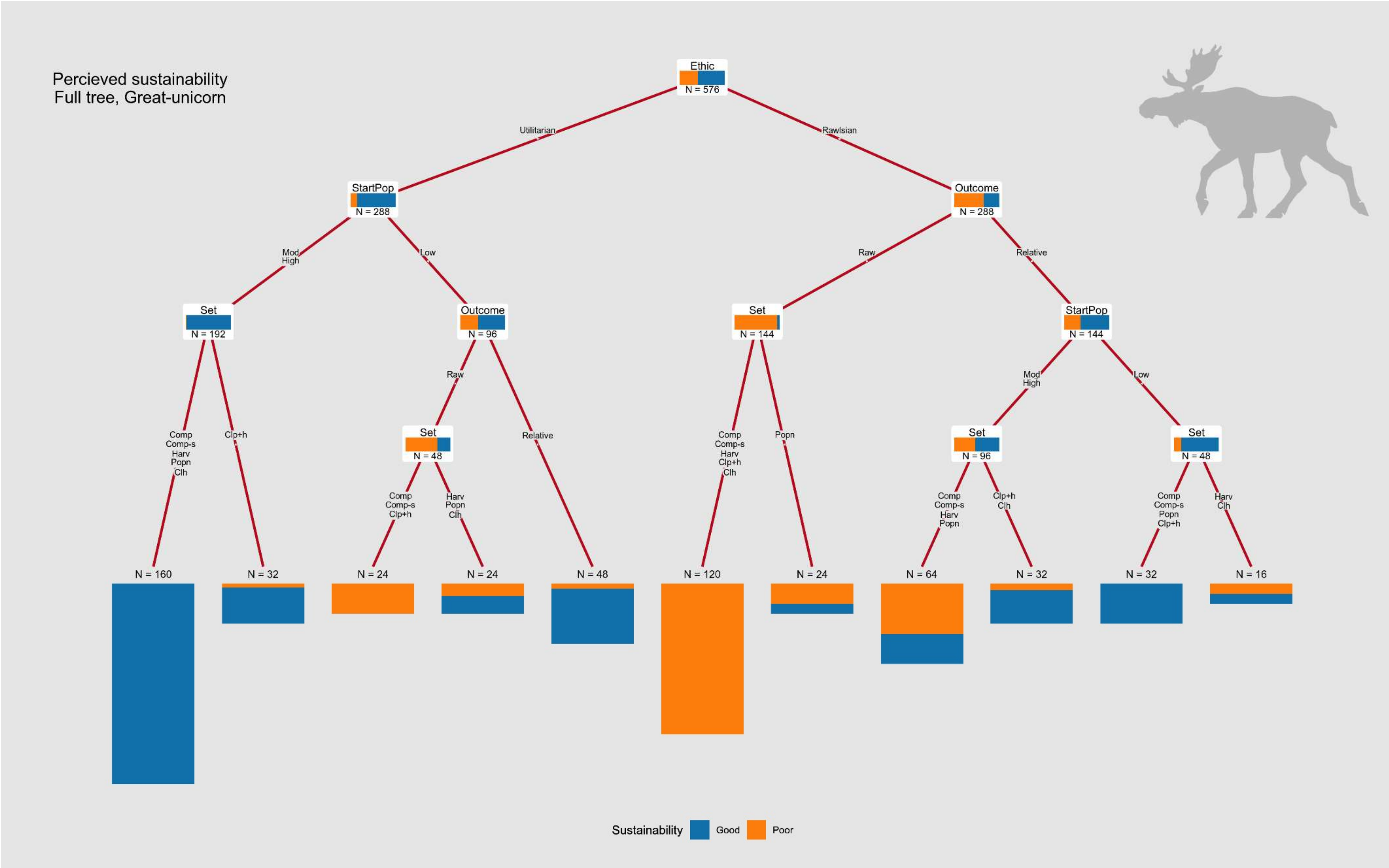
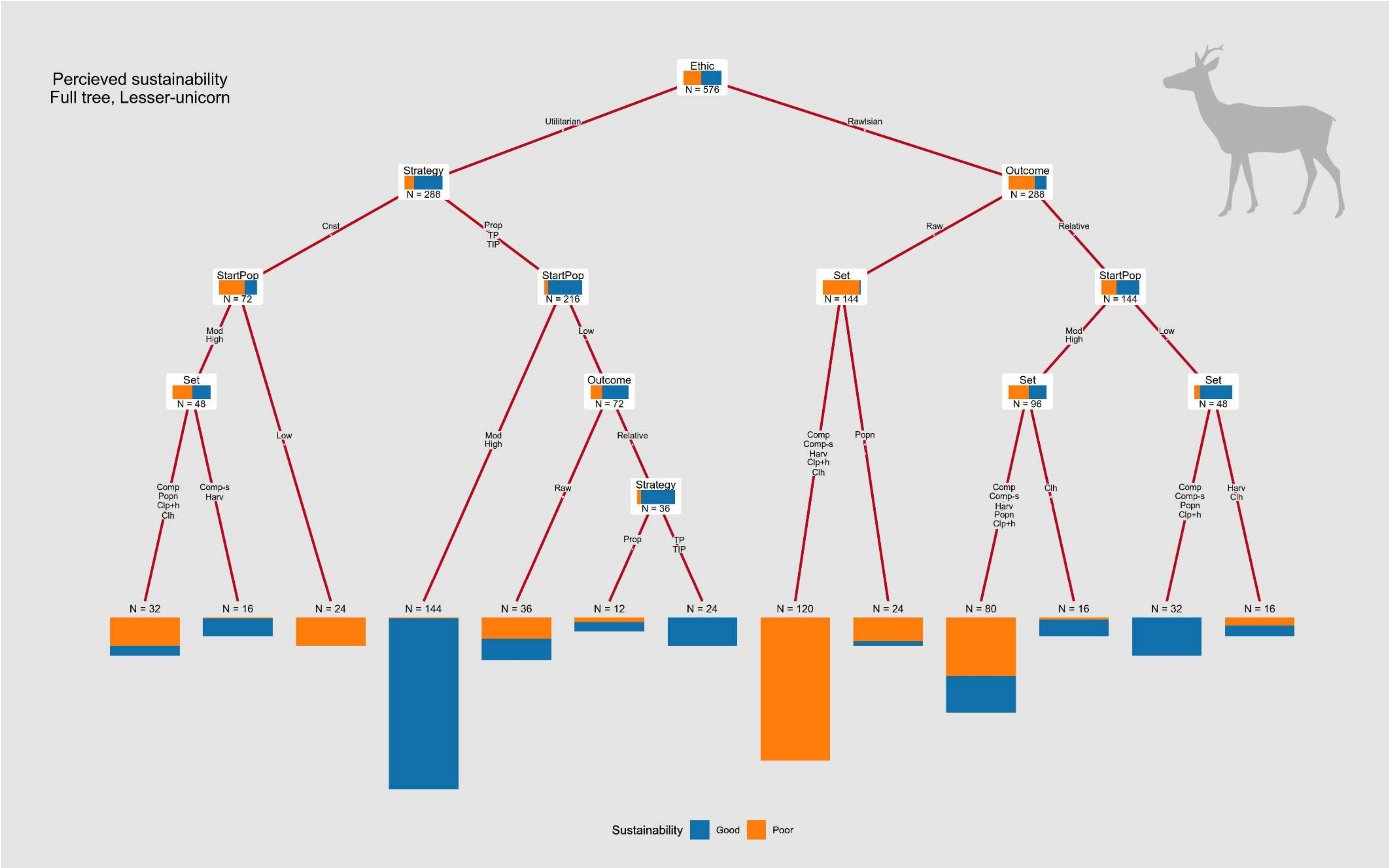


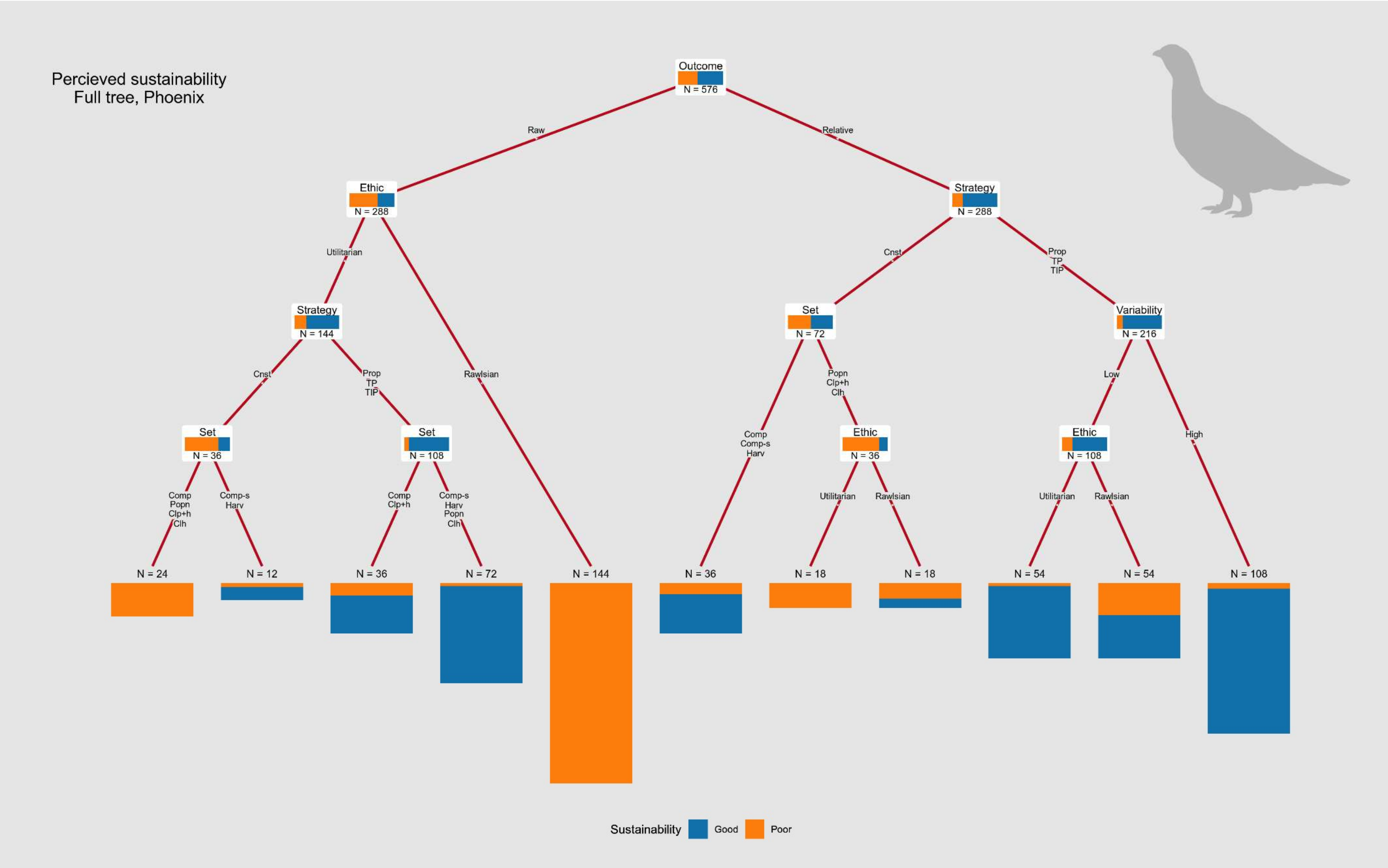
Figure S2.3.5 Perceived sustainability, Great-unicorn



## Figure S2.3.6 Perceived sustainability, Lesser-unicorn



1117 Figure S2.3.5 Perceived sustainability, Phoenix



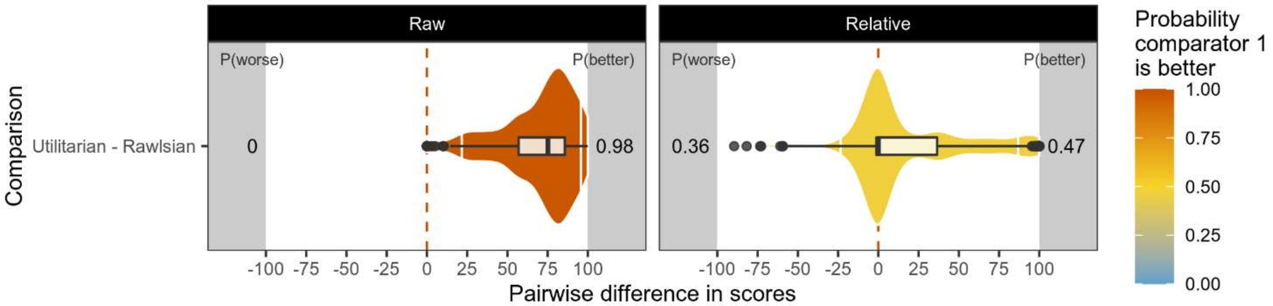

1118

## S2.4 Pairwise comparisons

As our framework systematically modelled all contextual factors in a factorial design, we are able to contrast factors based on the pairwise differences, i.e. when all other factors are held constant. This allows us to determine the independent influence of different contextual factors. Here we plot the distributions of the pairwise comparisons as comparator 1 – comparator 2, thus, when the score is negative, comparator 2 is better, and when the score is positive, comparator 1 is better.

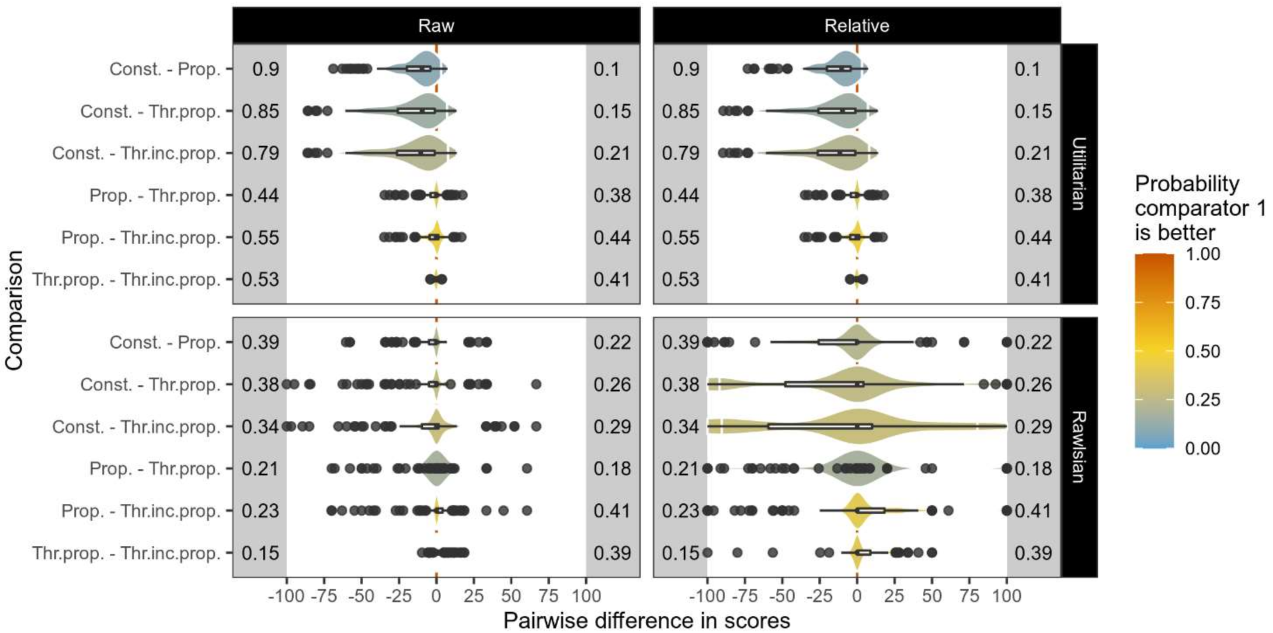
### Figure S2.4.1: Ethical perspective pairwise comparisons

Distributions of pairwise differences between ethical perspectives. Values in grey panels show proportions of 'better' and 'worse' outcomes for the comparisons. Note these may not sum to 1 as a percentage may not change. Violin plots are coloured by the median score. White lines within the violin plots mark the 5% and 95% quantiles, and the boxplots the median and quartiles, with whiskers extending to 1.5 times the interquartile range. There are n = 432 cases in each violin.



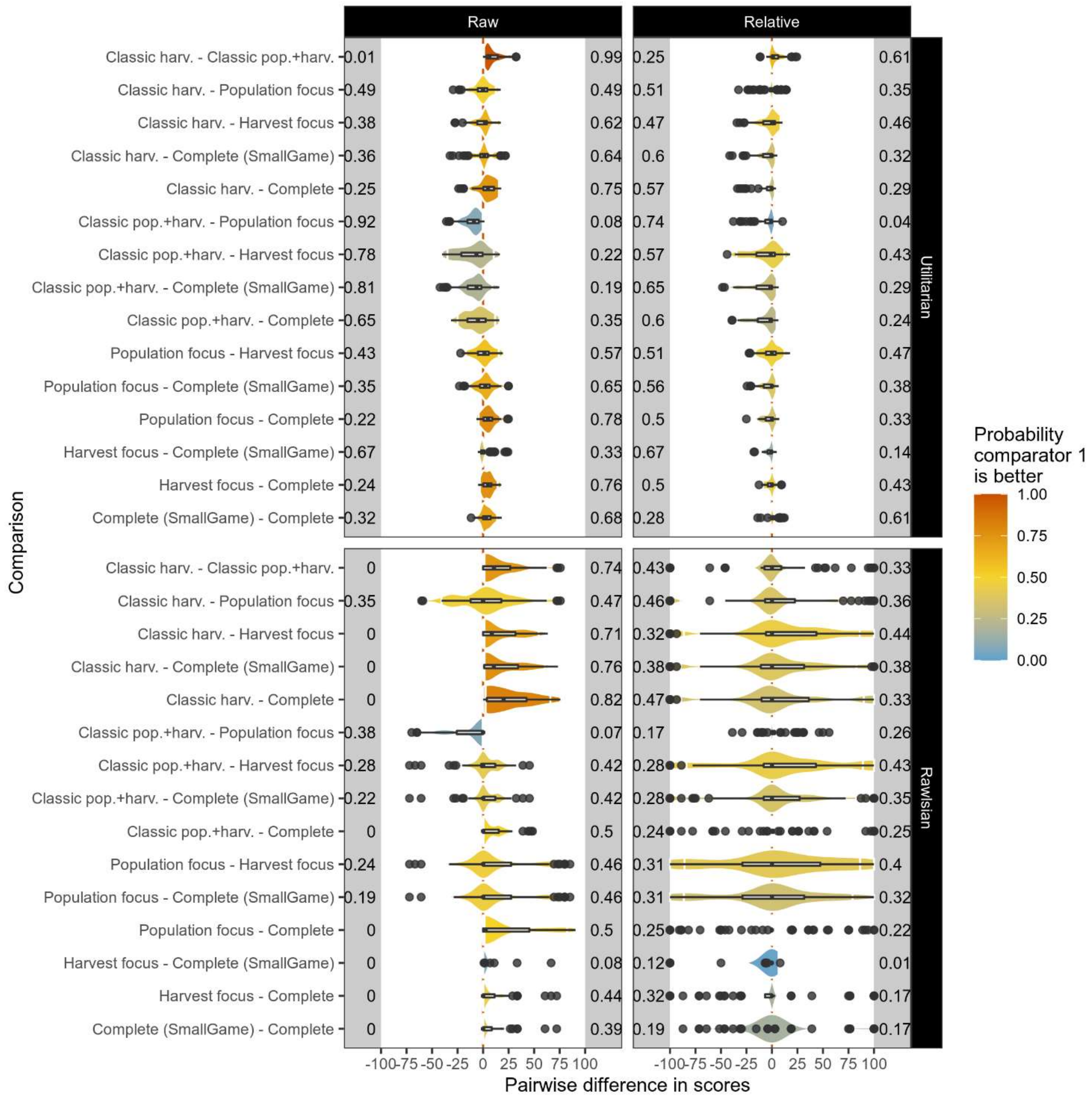

### Figure S2.4.2: Harvest strategy pairwise comparisons

Distributions of scores by, and pairwise differences between harvest strategies, differentiated by ethic and comparator. See Figure S2.2.1 for full description of plot details. There are n = 108 cases in each violin.



## Figure S2.4.3: Composite set pairwise comparisons

Distributions of scores by, and pairwise differences between composite sets, differentiated by ethic and comparator. See Figure S2.2.1 for full description of plot details. There are 72 cases in each violin.



## Figure S2.4.4: Environmental context comparisons

Distributions of scores by, and pairwise differences between, environmental context factors, differentiated by ethic and comparator. See Figure S2.2.1 for full description of plot details.