Appendix S1 A detailed description of the data sourcing and filtering of the population time series data for 1198 species, collection and rationalization for the extinction correlates, a full description of the statistical modelling methods employed, and the additional references cited in the Supplementary information.

Time-series data selection procedure

Time-series data were obtained from a variety of sources including compiled databases, primary scientific literature, books, and the internet. These included:

1. NERC Centre for Population Biology, Imperial College. *The Global Population Dynamics Database*. http://www.sw.ic.ac.uk/cpb/cpb/gpdd.html.

2. Patuxent Wildlife Research Center, USGS (USA), *Amphibian Count Database (ACD)*. http://www.mp2-pwrc.usgs.gov/cvs/ampCV/.

3. United States Fish and Wildlife Service (USFWS), *Waterfowl Population Status, 2002* (Appendix F. p.46-47). http://migratorybirds.fws.gov/reports/reports.html.

4. Cody, M.L. & Smallwood, I.A., eds. (1996). *Long-term Studies of Vertebrate Communities*. Academic Press, San Diego.

5. Primary literature found by searching for key terms (e.g. 'time-series') or known, well-studied species.

A full listing of all sources is given at the end of the supporting information, and descriptive statistics of the final database are given in Table S1.

The GPDD is the largest database of population time-series data currently available. It contains nearly 5000 separate time series for over 1400 animal and plant species. This database was used as our foundation data source, with additional time-series being added for any species not already present in the GPDD, or where superior time-series were available. For those intending to use resources such as the GPDD to make comparisons between species and across taxonomic groups, it is important that the data be brought into a common, consistent format. In its raw form, the data contained in the GPDD and other databases posed numerous difficulties for this, and many other sorts of quantitative investigation.

In its raw form, the data were not stored in a common, consistent format. To permit quantitative comparisons between species and across taxonomic groups, the following caveats must be considered:

1. Some species or taxonomic groups were well-studied and had many long-term time-series data available and so were over-represented in the GPDD compared to more poorly studied taxa.

2. Some time series gave the combined abundance of a number of different species within a loosely defined taxonomic group, e.g. deer, aphids.

3. Many single-species time series had poorly specified taxonomy such that they were only identifiable to the level of family or genus.

4. Study lengths varied from several months to 151 years.

5. Sampling frequencies included daily, monthly, seasonal, annual or even generational (with time intervals unspecified).

6. Measures of abundance varied across ten orders of magnitude.

7. Many time series contained zero abundance measures, which may have reflected extinction, extirpation, migration, cryptic behavior, inadequate sampling methods or insufficient sampling effort. The nature of zero values was usually unspecified.

8. Measures of abundance varied and included total population census, census of closed regional subpopulations, sampling within large open populations, counts of specific

demographic segments (eggs, juveniles, adults, calling males, nesting females), measures of density, indirect measures of species presence (scats, tracks), trapping rates, number of individuals harvested, total mass of harvests, and many other indices.

9. Some abundance measures included corrections for sampling intensity, but many time series were uncorrected and provided no means of accounting for this.

10. Many abundance measures were transformations of population size, including various ratios, logarithms and power relationships.

11. Sampling methodologies were diverse and difficult to classify without reference to the primary literature. Efforts by database managers to classify data type systematically, sampling units, and source units were clearly hampered by inconsistent reporting of such information in the source literature and the subjective nature of their classification system (leading to 373 different classifications for the 4933 time series in the GPDD).

12. Similarly, efforts by database managers to produce measures of the quality of each time series were also subjective.

To help address these issues, time series were accepted for analysis only if they met the following criteria:

1. Attributable to a single, identifiable species. Time series of organisms identifiable only to the genus or family level and those which were a combined abundance of several species were excluded.

2. If multiple abundance measures were provided within each year, then these were averaged to produce a single measure for each year.

3. All transformed abundance measures (mostly $\log_e$, $\log_{10}$ or $\log_e$+1 transformations) were back-transformed to produce values that were directly proportional to population size.

4. Where a single zero or sequence of zeros occurred in a time series, the first zero in each

run of zeros was converted to the lowest non-zero abundance measure for that time series so that it could be included as a low value for analysis (i.e. to permit the calculation of *r*). This assumes that single zeros are indicative of low abundance, below the threshold for observation. Preserving them as a low, non-zero value allows the pattern of the data to be preserved, whether it was a decline to a low-abundance state or recovery from it. Where there were two or more consecutive zeros, continued failure to observe the species may indicate migration, extirpation or inadequate sampling strategy. Given this uncertainty and the fact that successive zeros are uninformative for analyzing species' population dynamics, all remaining zeros in each run of zeros were treated as missing data. The sensitivity of our results to these zero-abundance observations is reported in the main text.

5. Time series were required to contain at least eight annual abundance transitions (at least nine years long if there were no years of missing data). For example, a ten year study with missing data for the sixth year would have been excluded because it contained only seven annual transitions.

6. Abundance measures in each time series were required to fluctuate over at least four different values. Time series exhibiting just two or three abundance states (due to low abundance or coarse measurement indices) were excluded.

7. To overcome the over-representation of a small number of well-studied species, only a single time series was allowed for each species. However, this produced the challenge of selecting a single 'best' time series from those available. This was achieved by a systematic quality scoring system designed to select the time series that gave an optimal balance of reliability, length and population size. Reliability scores were assigned subjectively to each time-series by the managers of the GPDD on a scale of 1 (harvest data for large with no indication of catch effort) to 5 (census data with accurate abundance values). Reliability scores were useful for identifying high-quality time series for each species, but many species had several time series with high reliability scores that varied substantially in

duration, and/or number of individuals surveyed. To obtain a balance of reliability score, time series length and size of the surveyed population, a quality index was calculated for each time series and only the highest-scoring time series was kept for each species.

Quality Score = [Reliability] + [TimeScore] + [PopScore] ÷ 2, where:

- [Reliability] = subjective reliability score given in GPDD and similarly to new data sets, with values ranging from 1 (poor) to 5 (excellent).
- [TimeScore] = standardize ($\log_e$ ([Timesteps] + [Changes])). Maximum and minimum values restricted to ± 2.5 standard deviations to reduce the influence of extreme outliers. Rescaled from 1 to 5.
- [Timesteps] = total number of transitions between consecutive annual abundance measures.
- [Changes] = number of consecutive annual changes in abundance. May be equal to [Timesteps] or less, if abundance is constant over consecutive years.
- [PopScore] = standardized ($\log_e$ [Average Population Size]). Maximum and minimum values restricted to ± 2.5 standard deviations to reduce the influence of extreme outliers. Rescaled from 1 to 5.

Extinction correlates

Following a review of the scientific literature on extinctions (especially Gilpin & Soulé 1986; McKinney 1997; Purvis *et al*. 2000 – see Table S3 for a complete listing), we decided on a set of nineteen morphological, life history, ecological and behavioral attributes that have been shown or postulated to correlate with extinction risk (see Table S2 and *Original Correlates*, below). Data were collected for 1198 species (the Microsoft Excel database is provided in supporting data). We then reduced in number and refined, based on theory and logic, the initial nineteen correlates to a set of six meaningful, derived variables (see Burnham & Anderson 2002) called composite predictors. These were:

P1, Threat Index [TI]: Species were considered to be *threatened* (score = 1) if they were either (1) legally protected in their native region, (2) listed under the IUCN Red List (IUCN 2005) as anything other than *least concern*, excluding *data deficient* and *not evaluated*, (3) where the global population numbered less than 5000 individuals, or (4) population size exhibits a continued decline. Otherwise deemed *lower risk* (score = 0). We also did another analysis in which just IUCN listing rather than TI was used as a predictor variable, but this made little difference to the fit of the saturated model (% deviance explained for GMVP changed from 0.24 % to 1.14 % when using IUCN instead of TI).

P2, Geographic Range [RA]: Range size or geographic distribution. Scores were assigned categorically and then converted to a continuous predictor from 0 to 1 (by subtracting the total by minimum score possible and then dividing by the range). A high score indicated an assumed high level of threat. Score criteria were: (1) global or continental range (migratory birds, weeds etc), (2) confined to a single phytogeographic or oceanic region, (3) confined to a single biome, (4) narrow endemic within 500 km^2 and (5) extremely narrow endemic within 50 km^2. The sensitivity of our results to the way in which range size was defined is reported in the main text.

P3, Human impact [HI]: This considered the extent of range or habitat loss as well direct loss (culling, etc.) or indirect loss (pollution, competition with weeds etc) and the severity of these losses. Scores for range decline were given according to the extent of loss, thus: (1) species occupies > 50 % of its former range, (2) species occupies 2 to 50 % of its former range and (3) species occupies < 2 % of its former range. Direct and/or indirect losses were scored by either 1 (where experienced) or 0 otherwise. Severity of losses was scaled as (1) slight, (2) moderate, (3) serious. Range decline and severity were given greater weighting by this method. Final score was converted to range from 0 to 1 by subtracting the minimum possible value from the end value and dividing by the range for each contributing metric, giving each element of the composite variable an equal contribution to the final value.

P4, Body size [BS]: Data were collected for all species as body length in millimeters. Body length or size was defined as tip-of-beak to tip-of-tail for birds, tip-of-snout to vent for reptiles and amphibians, tip-of-nose to tip-of-tail for mammals, tip-of-snout to tail-fork for fish, canopy height for plants, front of head to end of abdomen for insects, except Lepidoptera where wingspan was used. Measurements were converted using the natural logarithm. Body weight was available for vertebrate species, but not invertebrates or plants (669 species) and body length was used as the common measure.

P5, Ecological flexibility [EF]: Here we took dispersal ability, trophic level and the extent of ecological specialization to be surrogates of 'ecological flexibility', assuming that those species thought to be more 'flexible' than others are better adapted to change. Dispersal ability was defined as distance moved from place of origin in one generation. This logic allowed for relatively immobile taxa such as plants to be considered, given that seedling success is effectively parental success. A high score indicated an assumed high threat level. Dispersal ability was categorized as (1) > 10 000 km, (2) 1000 to 10 000 km, (3) 100 to 1000 km, (4) 10 to 100 km, (5) 1 to 10 km or (6) < 1 km. Trophic level was scored thus: (1) primary producer, (2) detritivore, (3) herbivore, (4) omnivore and (5) carnivore. Species at the top of the food web were assumed to be less flexible ecologically than those at the bottom (5). Ecological specialization or niche breadth considered synthetically both feeding specialization and habitat specialization (5). Scores were simply: (1) generalist and (2) specialist. Final score was converted to range from 0 to 1 by subtracting the minimum possible value from the end value and dividing by the range for each contributing metric, giving each element of the composite variable an equal contribution to the final value.

P6, Demographics [DE]: Considered as the reproductive life history of species (i.e. age at sexual maturity, fertility, reproductive strategy and longevity). Highly fecund, short-lived species were assumed to be more competitive than long-lived species with extended gestation periods, at least in response to short-term and major change related to human impacts. Fertility (number of eggs laid or young born per female per annum) was used because fecundity (number of young surviving to breeding age) was impossible to quantify in most cases. Generation length was discarded owing to inadequate data, but age at first breeding was used as an index of generation length for the simulation exercises. These parameters were categorized and additive values allowed for a final score (a high score indicated an assumed high level of threat). Categories were: Fertility (1) > 100 young per annum, (2) 5 to 100 young per annum and (3) < 5 young per annum. Longevity (maximum life span) was scored thus: (1) < 2 year, (2) 2 to 5 years and (3) > 5 years. Age at sexual maturity was scored: (1) < 1 year, (2) 1 to 3 years and (3) > 3 years. Reproductive strategy was simply (1) asexual, hermaphroditic or vegetative and (2) sexual. The final score was converted to range from 0 to 1 by subtracting the minimum possible value from the end value and dividing by the range for each contributing metric, giving each element of the composite variable an equal contribution to the final value.

An example of scores assigned (derived predictors and initial correlates) to three vertebrate species and two invertebrates is provided in Table S2.

Original Correlates

The initial nineteen correlates used to collect raw data were:

C1 Body weight: average adult weight (male and female) measured in grams to allow for small body size.

C2 Body length: average adult length (male and female) measured in mm. Rules mentioned above.

C3 Reproductive type: (1) asexual/vegetative/hermaphroditic or (2) sexual.

C4 Age at sexual maturity: average age at which an individual first mates given in months.

C5 Lifespan: maximum age attained by individuals in the wild measured in months.

C6 Generation length: average age of breeding adults at the time their young are born in months.

C7 Social grouping: taken to be a grouping of breeding adults. Categorized as (1) colonial, (2) gregarious, (3) small group, (4) pair and (5) solitary. Later considered nebulous and discarded.

C8 Dispersal ability: as mentioned in P5.

C9 Disturbance type: Direct (0 to 1) + Indirect (0 to 1).

C10 Fragmentation (range decline): as mentioned in P3.

C11 Geographic distribution: as mentioned in P2.

C12 Population size: categorical estimation of effective adult population at time of study. Categories were: (1) > 500 000, (2) 50 000 to 500 000, (3) 5000 to 50 000, (4) 500 to 5000, (5) 50 to 500 and (6) < 50 breeding individuals.

C13 Fertility: as mentioned in P6.

C14 Population trend: trend at time of study, given as (1) increasing (2) stable and (3) declining.

C15 Trophic level: as mentioned in P5.

C16 Niche breadth: (1) generalist or (2) specialist. Further mention given in P5.

C17 Relationship with Homo sapiens*:* scored as (1) positive or (2) negative according to whether there was any apparent benefit gained through people, such as introduced weeds or agricultural crops.

C18 IUCN listing: Following IUCN (see IUCN 2005), categories were extinct (EX), extinct in the wild (EW), critically endangered (CR), endangered (EN), vulnerable (VU), near threatened (NT), least concern (LC), data deficient (DD) or not evaluated (NE).

C19 Legal protection: categorized according to whether a species was either protected (2) in native region, or listed under CITES Appendices I or II. If unprotected assigned value of 1.

A number of attributes listed within the literature as being correlated with extinction risk were excluded from the outset because reliable and consistent data for these measures were not available for most species. These extinction correlates were:

1. *Biome*: Temperate, Tropical, Arid, Freshwater and Marine. Later considered to be too unselective and difficult to combine with other predictors (being categorical, required the estimation of an extra four parameters in the GLMMs).

2. *Home range:* Data on *average home range size* was difficult to obtain from the literature (particularly for birds, fish, amphibians, reptiles and invertebrates). Home range size data for mammals was questionable at best.

3. *Behavior*: Courtship behavior, social behavior, inter-specific interactions and habitat selection. These variables were too difficult to source and quantify, especially given time constraints.

4. *Physiology*: Metabolic efficiency and disease resistance. Data were mostly unavailable.

5. *Environment*: Abundance (density) of resources, abundance of interacting species. Data were mostly unavailable.

6. *Population structure and fitness*: Age structure, sex ratio, saturation density. Data were mostly unavailable.

References

Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Multimodal Inference: A Practical Information-Theoretic Approach*. 2nd edn. Springer-Verlag, New York.

Gilpin, M.E. & Soulé, M.E. (1986). Minimum viable populations: processes of species extinction. In: *Conservation Biology: The Science of Scarcity and Diversity* (ed. Soulé, M.E.). Sinauer, Sunderland, MA, pp. 19-34.

IUCN (2005). *IUCN Red List of Threatened Species*. World Conservation Union (www.redlist.org).

McKinney, M.L. (1997). Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annu. Rev. Ecol. Syst.*, 28, 495-516.

Purvis, A., Gittleman, J.L., Cowlishaw, G. & Mace, G.M. (2000). Predicting extinction risk in declining species. *P. Roy. Soc. Lond. B. Bio.*, 267, 1947-1952.

Table S1 Summary statistics (averages across groups, by taxon) of the population dynamics time-series dataset of 1198 species, where n = number of species, q = number of annual observed transitions in population size, r = average rate of population change (log annual ratio of successive densities), σ^2_r = variance of r, CVN = coefficient of variation in population size.

Taxonomic group	n	q	r	σ^2_r	CVN
All species	1198	22	0.005	0.733	0.610
Invertebrates	639	19	0.010	0.719	0.604
Vertebrates	529	26	0.000	0.680	0.606
Plants	30	13	-0.014	1.940	0.817
Marine Invertebrates	35	19	0.019	1.155	0.833
Insects	604	18	0.009	0.694	0.591
Fish	115	21	-0.001	0.808	0.756
Herpatiles	37	14	-0.034	1.121	0.731
Birds	225	27	0.002	0.412	0.497
Mammals	152	31	0.005	0.873	0.625

Table S2 An example of scores assigned to three vertebrate species (mammal, bird and fish) and two invertebrates (moth and aphid).

Species / Predictors	*Panthera leo* Lion	*Thunnus albacares* yellowfin tuna	*Haliaeetus leucocephalus* bald eagle	*Horisme tersata* fern moth	*Sitobion avenae* grain aphid
P1 Conservation	1	1	1	0	0
P2 Geographic range	0	0	0.25	0.25	0
P3 Human impact	0.8	0.4	0.6	0.2	0
P4 Body size	7.882	7.090	7.601	3.497	0.6931
P5 Ecological	0.524	0.286	0.405	0.357	0.357
P6 Demographics	0.833	0.667	0.833	0.167	0.167
C1 Body weight	163353	40000	4740		
C2 Body length	2650	1200	2000	33	2
C3 Reproductive type	2	2	2	2	1
C4 Sexual maturity	42	24	48	8	1
C5 Lifespan	168	144	336	10	9
C6 Generation length	40	48	48	8	1
C7 Grouping	2	1	3	5	2
C8 Dispersal	3	1	2	4	4
C9 Disturbance	1 + 1 + 3	1 + 1 + 2	1 + 1 + 1	0 + 1 + 1	1 + 0 + 1
C10 Range decline	2	1	2	1	1
C11 Distribution	1	1	2	2	1
C12 Population size	3	1	3	1	1
C13 Fertility	1.5	2000000	2	550	85
C14 Population trend	3	3	2	2	2
C15 Trophic level	4	5	5	3	3
C16 Niche breadth	1	1	2	1	2
C17 Relation with	2	2	2	2	1
C18 IUCN listing	VU	LC	CD	LC	LC
C19 Legal protection	2	1	2	1	1