

## Reductions in global biodiversity loss predicted from conservation spending

**Authors:** Anthony Waldron<sup>1,2</sup>, Daniel C. Miller<sup>2</sup>, Dave Redding<sup>3</sup>, Arne Mooers<sup>4</sup>, Tyler S. Kuhn<sup>5</sup>, Nate Nibbelink<sup>6</sup>, J. Timmons Roberts<sup>7</sup>, Joseph A. Tobias<sup>1,8</sup> & John L. Gittleman<sup>9</sup>

### **Affiliations:**

1. Edward Grey Institute, Department of Zoology, Oxford University, Oxford OX1 3PS, UK.
2. Department of Natural Resources and Environmental Science, University of Illinois, Urbana-Champaign, IL61801, USA.
3. Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK.
4. Biology Department, Simon Fraser University, Burnaby, BC V5A 1S6, Canada.
5. Scimitar Scientific, Whitehorse, YT, Y1A 6V6, Canada.
6. Warnell School of Forestry & Natural Resources, University of Georgia, Athens, GA 30602, USA.
7. Institute for the Study of Environment and Society, Brown University, Providence, RI 02912, USA.
8. Department of Life Sciences, Imperial College London, Silwood Park, Buckhurst Road, Ascot, Berkshire, SL5 7PY, UK.
9. Odum School of Ecology, University of Georgia, Athens, GA, 30602, USA.

Correspondence and requests for materials should be addressed to [anthony.waldron@zoo.ox.ac.uk](mailto:anthony.waldron@zoo.ox.ac.uk) / [anthonywaldron@hotmail.com](mailto:anthonywaldron@hotmail.com)

**Keywords:** Conservation, Aichi targets, Sustainable Development Goals, Convention on Biological Diversity, conservation finance, conservation impact, conservation effectiveness, economic development, IUCN Red List, extinction

**Halting global biodiversity loss is central to both the Convention on Biological Diversity (CBD) and United Nations Sustainable Development Goals (SDGs)<sup>1,2</sup>, but success to date has been very limited<sup>3-5</sup>. A critical determinant of overall strategic success (or failure) is the financing committed to biodiversity<sup>6-9</sup>; however, financing decisions are still hindered by considerable uncertainty over what any investment is likely to achieve<sup>6-9</sup>. For greater effectiveness, we need an evidence-based model (EBM)<sup>10-12</sup> showing how conservation spending quantitatively reduces the rate of loss. Here, we empirically quantify how i\$14.4 billion of conservation investment reduced biodiversity loss across 109 signatory countries between 1996 and 2008, by an average 29% per country. We also show that biodiversity change in signatory countries can be predicted with high accuracy, using a dual model that combines the positive impact of conservation investment with the negative impact of economic, agricultural and population growth (i.e. human development pressures)<sup>13-18</sup>. Decision-makers can use this dual model to forecast the improvement that any proposed biodiversity budget would achieve under various scenarios of human development pressure, comparing those forecasts to any chosen policy target (including the CBD and SDGs). Importantly, we further find that spending impacts shrink as human development pressures grow, implying that funding may need to increase over time. The model therefore offers a flexible tool for balancing the SDGs of human development and biodiversity, by predicting the dynamic changes needed in conservation finance as human development proceeds.**

49  
50 The rapid loss of global biodiversity has major consequences for human wellbeing<sup>5,19</sup> and so  
51 governments worldwide have committed reducing those losses through multiple international  
52 agreements, including the CBD and SDG frameworks<sup>1,2</sup>. However, strategic outcomes to date  
53 have been poor: we missed the 2010 CBD target and now seem likely to also miss the 2020  
54 Aichi biodiversity targets<sup>3,4</sup>. As outlined in Aichi target 20 and SDG17, one of the most  
55 important determinants of policy success is our ability to correctly decide (and secure) the level  
56 of financing needed to resource overall biodiversity-conservation strategies<sup>1,2,6-8</sup>. A second key  
57 way to improve on currently poor outcomes is to take a more evidence-based approach, in which  
58 decision making is guided by reliable evaluations of past successes and failures (“conservation  
59 impact assessments”)<sup>10-12</sup>. In many fields, the financing of strategic goals is fundamentally  
60 evidence-based, analysing previous spending outcomes to guide current budget decisions<sup>20,21</sup>.  
61 Surprisingly, however, no study has yet tested whether global conservation investment has  
62 actually reduced biodiversity decline across CBD signatory countries, nor quantified the  
63 differential impacts of different funding levels.

64  
65 A second key policy need is for models that reliably predict biodiversity decline, so that future  
66 losses can be forecast and timely action taken<sup>15,22</sup> (as already occurs with climate change<sup>23</sup>). In  
67 bio-political science, predictive models typically quantify how biodiversity loss is driven by  
68 human socioeconomic pressures, such as economic or agricultural expansion<sup>14-16,24</sup>. To date,  
69 conservation impact assessments and predictive decline models have largely developed as  
70 separate major fields, despite their outcomes being strongly interdependent. It is rarely possible  
71 to accurately measure the impact of one factor (either spending or pressures) on biodiversity  
72 without accounting for the influence of the other factor<sup>3,25</sup>. To make accurate predictions for  
73 policy use, we therefore need unified models that treat biodiversity change as the simultaneous  
74 outcome of pressures and their impact, plus conservation and its impact (henceforth, “pressures-  
75 and-conservation-impact (PACI) models”). Indeed, one of the core challenges for the SDGs is to  
76 balance (or trade off) the often-conflicting goals of human development (e.g. SDGs 1, 2 & 8) and  
77 biodiversity conservation (SDG 15)<sup>2,14-18,24</sup>. To measure this trade-off, policymakers need  
78 models that unite these two aspects. Finally, such models need to apply to the key geopolitical  
79 decision-making scale for the CBD and SDGs – sovereign countries – demanding finer  
80 geographic resolution than common planet-scale approaches<sup>3,7</sup>.

81  
82 Here, we use empirical evidence to develop a unified PACI model at the sovereign country scale,  
83 by statistically quantifying how changing human pressures drive biodiversity decline while  
84 conservation spending reduces it. As such, the model informs policymakers not just what  
85 biodiversity losses to expect but more constructively, how changes in conservation resourcing  
86 can reduce those expected losses<sup>3</sup>. We also show how the impacts of spending and pressures  
87 depend predictably upon national socioeconomic contexts, and thus how they may change over  
88 time.

89  
90 A standard policy measure of biodiversity change (usually, decline) is the planet-scale sum of all  
91 changes in individual species’ IUCN Red List status, using well-known taxa as a proxy for  
92 biodiversity<sup>3,26</sup>. To calculate biodiversity change at the decision-making scale of sovereign  
93 signatory countries (hereafter each country’s “biodiversity decline score” or BDS), we took Red  
94 List status changes for all global bird and mammal species for 1996–2008 (see Methods for

95 justification and details) and portioned them out among all countries where each species is found  
96 (treating the few status improvements as negative fractions). We then summed all decline  
97 fractions for each country to calculate BDSs<sup>8,26</sup> (Figure 1, Supplementary Table 1). It is  
98 noteworthy that 60% of total BDS for the globe was found in only seven countries: Indonesia,  
99 Malaysia, Papua New Guinea, China, India, Australia, and the USA (principally Hawai'i). Seven  
100 countries had net biodiversity improvements (negative BDSs): Mauritius, Seychelles, Fiji,  
101 Samoa, Tonga, Poland and Ukraine. (See Extended Data Figure 1 for average BDS per species).

102  
103 To be useful in policymaking, models of biodiversity change need to have simple generality and  
104 demonstrated forecasting accuracy. Therefore, we first built PACI regression models to predict  
105 known BDS, using national-level data on strict-sense conservation spending (annualised, see  
106 Methods) plus the broad socioeconomic pressures of GDP growth, agricultural expansion (and  
107 its relationship to forest loss), human population growth, and changing governance quality  
108 (Extended Data Table 1, Supplementary Table 2). We then tested forecasting accuracy by using  
109 cross-validation, which repeatedly presents the model with data it has not seen and asks it to  
110 predict a known outcome (see Methods). BDS data were continuous zero-inflated due to multiple  
111 species-poor countries with no status changes, so we used two-part models<sup>27</sup> in which the  
112 “continuous” part (n=50) models BDS after truncating the long tail of zeroes, and the “binomial”  
113 part (n=109) models whether BDS is zero or non-zero across all countries. We tested for context  
114 dependence by fitting several hypothesized interactions (Methods, Extended Data Table 1).

115  
116 In the best-fitting regression models (Table 1), we found that conservation spending strongly  
117 reduced decline (i.e. BDS, Figure 2), whereas GDP growth and agricultural expansion tended to  
118 increase it (Figure 3). Although forest loss was often significant, the best-fitting predictive model  
119 favoured more generalized terms (Table 1, Supplementary Discussion). Interaction terms  
120 revealed several context-dependent nuances (see Supplementary Discussion). The GDP growth  
121 effect decreased as baseline GDP decreased, becoming non-significant in the poorest countries  
122 (Figure 3). Agricultural expansion had a deleterious impact in countries with relatively low  
123 percentages of land devoted to agriculture (such as Malaysia and Peru), but was not statistically  
124 significant in countries with mid-to-high percentages such as Bangladesh (Figure 3). The  
125 binomial part also suggested that the impact of agricultural expansion could be greatly reduced  
126 by improvements in the quality of national governance (Extended Data Figure 2), and that the  
127 deleterious impact of GDP became stronger as human population growth increased, i.e. the  
128 combined impact of two pressures was greater than the sum of its parts (Table 1). Finally,  
129 conservation spending was more effective in poorer countries than in higher-income ones, and  
130 spending also had a greater impact when more species were threatened in the first place  
131 (Extended Data Figure 3).

132  
133 Both model parts accurately predicted historical declines ( $R^2 = 0.85$  in the continuous part;  
134 accuracy = 94% in binomial part; Extended Data Figure 4) and were robust to several sensitivity  
135 tests (Supplementary Results, Extended Data Table 4). They also had high forecasting accuracy  
136 in cross-validation (82% continuous part; 85% binomial part). Our PACI models therefore have  
137 immediate application to several major policy needs. They can predict not only future  
138 biodiversity declines<sup>15,22</sup>, but also how changes to a key policy instrument – the high-level  
139 financial resourcing of biodiversity conservation – will quantifiably reduce the declines  
140 expected. To illustrate this feature, we used the model to predict the impact of spending an extra

141 i\$5 million in each country (such that the overall global annual budget was increased by 42%,  
142 Supplementary Table 3). Outcomes for all countries are shown in Supplementary Table 3 (see  
143 also Figure 1) but to give an example: in the mega-diverse countries of PNG and Peru, the model  
144 predicted reductions in decline (BDS) of 33% and 54% respectively. We also used the model to  
145 back-predict how much biodiversity loss was prevented by post-Earth Summit conservation  
146 financing<sup>8,28</sup>, estimating that on average (median), losses per country were 29% less than would  
147 otherwise have occurred (Methods).

148  
149 The model could also be used to predict the funding each country needs to achieve specific  
150 biodiversity policy goals, including the CBD and SDG targets. Importantly, however, our results  
151 demonstrate how the cost of meeting any target constantly changes as the levels of  
152 socioeconomic pressure change. For example, if Peru had wanted to achieve 50% less decline by  
153 2008, then with pressures at their 1992-2003 levels, the model predicts that an extra \$4.6m  
154 annually would have been needed annually. However, at current (2001-2012 mean) levels of  
155 pressure, that figure would rise to \$5.7m (constant international dollars). Our model explicitly  
156 accounts for these changes in socioeconomic context, and so an appropriate policy use would be  
157 to take various scenarios of economic, agricultural and population change, and then predict  
158 biodiversity outcomes at different funding levels for each scenario, comparing them to targets. In  
159 particular, the model can be used to help resolve problems of discordance between the SDGs for  
160 biodiversity and human development, by quantifying how any negative effects of economic and  
161 agricultural growth can be balanced out by short-term increases in conservation funding (thereby  
162 creating a breathing space to develop more sustainable pathways to national growth<sup>18</sup>.)

163  
164 We caution that an unmeasured variable correlated with conservation spending could  
165 conceivably explain some of the spending impact; that the co-benefits of spending for taxa other  
166 than birds and mammals remain unknown; that species declines too small to affect Red List  
167 status will not be accurately predicted and will require different approaches<sup>29</sup>; and that long-  
168 distance effects such as Chinese demand for African ivory<sup>30</sup> were beyond the scope of our  
169 model. However, our general PACI approach should be flexible enough to accommodate such  
170 additions in the future.

171  
172 At a time when the outlook for biodiversity often seems very bleak<sup>4,5</sup>, our results present a  
173 constructive opportunity for global biodiversity policy, showing how increases in conservation  
174 investment can lead to major, quantifiable improvements. However, set against this note of  
175 optimism, our model also underlines how conservation spending may need to constantly increase  
176 (or evolve) to counterbalance the continuing intensification of human development  
177 pressures<sup>5,18,24</sup>. By empirically demonstrating how limited levels of investment have already led  
178 to a partial reduction in biodiversity loss, our findings may ultimately encourage decision-makers  
179 to commit the full finance needed<sup>7</sup> to significantly reduce or halt global losses, in line with our  
180 CBD and SDG commitments<sup>1,2</sup>.

181  
182  
183 **Acknowledgments:** This research was supported by UKDWP (A.W.), the USDA National  
184 Institute of Food and Agriculture Hatch project 1009327 (A.W. and D.C.M.), the MacArthur  
185 Foundation through the Advancing Conservation in a Social Context research initiative (D.C.M.  
186 and J.T.R.), Natural Sciences and Engineering Research Council Canada Discovery and

187 Accelerator Grants (A.O.M.), the UK Natural Environment Research Council (J.A.T.) and the  
188 Odum School of Ecology (J.L.G.). We thank J. Drake, P. Holland and P. Stephens and four  
189 anonymous referees for comments on earlier manuscripts. Detailed methods, additional results  
190 and data are available in the online supplementary material.

191  
192 **Author Contributions:** A.W. conceived the study and analysed the data, based on ideas from  
193 D.C.M, A.O.M. and J.L.G.; A.W., D.C.M., D.R., N.N. and J.T.R. collected the data; A.W.,  
194 A.O.M., T.S.K., D.C.M, J.L.G. and J.A.T. wrote the paper with contributions from all other  
195 authors.

196  
197 **Author Statement:** Reprints and permissions information is available at  
198 [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Correspondence  
199 and requests for materials should be addressed to anthonywaldron@hotmail.com.

200  
201  
202

203

204 **REFERENCES**

- 205 1. CBD. Conference of the Parties Decision X/2: Strategic plan for biodiversity 2011-2020.  
206 (2010). Available at: [www.cbd.int/decision/cop?id=12268](http://www.cbd.int/decision/cop?id=12268).
- 207 2. UN. *Transforming our world: the 2030 agenda for sustainable development*. (2015).
- 208 3. Hoffmann, M. *et al.* The impact of conservation on the status of the world's vertebrates.  
209 *Science* **330**, 1503–1509 (2012).
- 210 4. Tittensor, D. P. *et al.* A mid-term analysis of progress toward international biodiversity  
211 targets. *Science* **346**, 241–244 (2014).
- 212 5. Pimm, S. L. *et al.* The biodiversity of species and their rates of extinction, distribution,  
213 and protection. *Science* **344**, 1246752 (2014).
- 214 6. James, A. N., Gaston, K. J. & Balmford, A. Balancing the Earth's accounts. *Nature* **401**,  
215 323–324 (1999).
- 216 7. McCarthy, D. P. *et al.* Financial costs of meeting global biodiversity conservation targets:  
217 current spending and unmet needs. *Science* **338**, 946–949 (2012).
- 218 8. Waldron, A. *et al.* Targeting global conservation funding to limit immediate biodiversity  
219 declines. *PNAS* **110**, 12144–12148 (2013).
- 220 9. UNEP/CBD. *Decisions adopted by the conference of the parties to the Convention on*  
221 *Biological Diversity at its 11th meeting*. (2013).
- 222 10. Sutherland, W. J., Pullin, A. S., Dolman, P. M. & Knight, T. M. The need for evidence-  
223 based conservation. *Trends Ecol. Evol.* **19**, 305–8 (2004).
- 224 11. McKinnon, M., Cheng, S. H., Garside, R., Masuda, Y. J. & Miller, D. C. Sustainability:  
225 map the evidence. *Nature* **528**, 185–187 (2015).
- 226 12. Miteva, D., Pattanayak, S. & Ferraro, P. J. Analysis of biodiversity policy instruments:  
227 what works, and what doesn't. *Oxford Rev. Econ. Policy* **28**, 69–92 (2012).
- 228 13. Cardillo, M. *et al.* The predictability of extinction: biological and external correlates of  
229 decline in mammals. *Proc. Biol. Sci.* **275**, 1441–8 (2008).
- 230 14. Naidoo, R. & Adamowicz, W. L. Effects of economic prosperity on numbers of threatened  
231 species. *Conserv. Biol.* **15**, 1021–1029 (2001).
- 232 15. Freytag, A., Vietze, C. & Volkl, W. What drives biodiversity? An empirical assessment of  
233 the relation between biodiversity and the economy. *Int. J. Ecol. Econ. Stat.* **24**, 1–16  
234 (2012).
- 235 16. Smith, R., Muir, R., Walpole, M., Balmford, A. & Leader-Williams, N. Governance and  
236 the loss of biodiversity. *Nature* **426**, 67–70 (2003).
- 237 17. Roe, D., Elliott, J., Sandbrook, C. & Walpole, M. *Biodiversity conservation and poverty*  
238 *alleviation*. (Wiley-Blackwell, 2013).
- 239 18. Adams, W., Aveling, R. & Brockington, D. Biodiversity conservation and the eradication  
240 of poverty. *Science* **306**, 1146–1149 (2004).
- 241 19. Cardinale, B. J. *et al.* Biodiversity loss and its impact on humanity. *Nature* **486**, 59–67  
242 (2012).
- 243 20. Schwartländer, B. *et al.* Towards an improved investment approach for an effective  
244 response to HIV/AIDS. *Lancet* **377**, 2031–2041 (2011).
- 245 21. HM Treasury. Public sector business cases: using the five case model. Green Book  
246 Supplementary Guidance on Delivering Public Value from Spending Proposals. *HM*  
247 *Treasury* 1–152 (2013). doi:10.1007/s13398-014-0173-7.2
- 248 22. Collen, B. & Nicholson, E. Taking the measure of change. *Science* **346**, 166–167 (2014).

- 249 23. Stern, N. *The economics of climate change: the Stern review*. (Cambridge University  
250 Press, 2007).
- 251 24. Maxwell, S. L., Fuller, R. A., Brooks, T. M. & Watson, J. E. Biodiversity: the ravages of  
252 guns, nets and bulldozers. *Nature* **536**, 143–145 (2016).
- 253 25. Ferraro, P. J. *et al.* Estimating the impacts of conservation on ecosystem services and  
254 poverty by integrating modeling and evaluation. *Proc. Natl. Acad. Sci.* **112**, 7420–7425  
255 (2015).
- 256 26. Rodrigues, A. S. L. *et al.* Spatially Explicit Trends in the Global Conservation Status of  
257 Vertebrates. *PLoS One* **9**, e113934 (2014).
- 258 27. Basu, A. & Manning WG. Issues for the next generation of health care cost analyses. *Med*  
259 *Care* **47**, S109-14 (2009).
- 260 28. Miller, D. C., Agrawal, A. & Timmons Roberts, J. Biodiversity, governance, and the  
261 allocation of international aid for conservation. *Conserv. Lett.* **6**, 12–20 (2013).
- 262 29. Donald, P. F. *et al.* International conservation delivers benefits for birds in Europe.  
263 *Science* **317**, 810–813 (2007).
- 264 30. Vandegrift, J. Elephant poaching: CITES failure to combat the growth in Chinese demand  
265 for ivory. *Virginia Environ. Law J.* **31**, 102 (2013).
- 266  
267  
268
- 269 31. Butchart, S. H. M. *et al.* Measuring global trends in the status of biodiversity: Red list  
270 indices for birds. *PLoS Biol.* **2**, e383 (2004).
- 271 32. Butchart, S. H. M., Akcakaya, H. R., Kennedy, E. & Hilton-Taylor, C. Biodiversity  
272 indicators based on trends in conservation status: Strengths of the IUCN red list index.  
273 *Conserv. Biol.* **20**, 579–581 (2006).
- 274 33. UNSD. Sustainable Development Goals indicators. (2016). Available at:  
275 <http://unstats.un.org/sdgs/>. (Accessed: 15th October 2016)
- 276 34. Metrick, A. & Weitzmann, M. L. Conflicts and choices in biodiversity preservation. *J.*  
277 *Econ. Perspect.* **12**, 21–34 (1998).
- 278 35. Butchart, S. H. M. *et al.* Using Red List Indices to measure progress towards the 2010  
279 target and beyond. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **360**, 255–268 (2005).
- 280 36. Channell, R. & Lomolino, M. V. Dynamic biogeography and conservation of endangered  
281 species. *Nature* **403**, 84–86 (2000).
- 282 37. Hastie, T., Tibshirani, R. & Friedman, J. *The elements of statistical learning: data mining,*  
283 *inference, and prediction*. (Springer, 2009).
- 284 38. Waldron, A. *et al.* Turkey’s biodiversity funding on the rise. *Science* **341**, 1173 (2013).
- 285 39. International Bank for Reconstruction and Development and World Bank. Global  
286 Purchasing Power Parities and Real Expenditures. 2005 International Comparison  
287 Programme. (2008).
- 288 40. Miller, D. C. Explaining global patterns of international aid for linked biodiversity  
289 conservation and development. *World Dev.* **59**, 341–359 (2014).
- 290 41. Barrett, C. B., Gibson, C. C., Hoffman, B. & McCubbins, M. D. The complex links  
291 between governance and biodiversity. *Conserv. Biol.* **20**, 1358–66 (2006).
- 292 42. McNeely, J. A. Conserving forest biodiversity in times of violent conflict. *Oryx* **37**, 142–  
293 152 (2003).
- 294 43. Hanson, T. *et al.* Warfare in Biodiversity Hotspots. *Conserv. Biol.* **23**, 578–587 (2009).

- 295 44. Kaufmann, D., Kraay, A. & Mastruzzi, M. *The Worldwide Governance Indicators*  
296 *Methodology and Analytical Issues*. (2010). Available at [www.govindicators.org](http://www.govindicators.org).
- 297 45. Mills, J. H. & Waite, T. A. Economic prosperity, biodiversity conservation, and the  
298 environmental Kuznets curve. *Ecol. Econ.* **68**, 2087–2095 (2009).
- 299 46. Dietz, S. & Adger, W. N. Economic growth, biodiversity loss and conservation effort. *J.*  
300 *Environ. Manage.* **68**, 23–35 (2003).
- 301 47. Asafu-Adjaye, J. Biodiversity loss and economic growth: a cross-country analysis.  
302 *Contemp. Econ. Policy* **21**, 173–185 (2003).
- 303 48. Donald, P. F. Biodiversity impacts of some agricultural commodity production systems.  
304 *Conserv. Biol.* **18**, 17–37 (2004).
- 305 49. Cardillo, M. *et al.* Human population density and extinction risk in the world’s carnivores.  
306 *PLoS Biol.* **2**, E197 (2004).
- 307 50. McPherson, M. A. & Nieswiadomy, M. L. Environmental Kuznets curve: threatened  
308 species and spatial effects. *Ecol. Econ.* **55**, 395–407 (2005).
- 309 51. Majumder, P., Berrens, R. P. & Bohara, A. K. Is There an Environmental Kuznets Curve  
310 for the Risk of Biodiversity Loss? *J. Dev. Areas* **39**, 175–190 (2006).
- 311 52. Pandit, R. & Laband, D. N. Economic well-being, the distribution of income and species  
312 imperilment. *Biodivers. Conserv.* **18**, 3219–3233 (2009).
- 313 53. Kerr, J. T. & Currie, D. Effects of human activity on global extinction risk. *Conserv. Biol.*  
314 **9**, 1528–1538 (1995).
- 315 54. Meyfroidt, P. & Lambin, E. Global forest transition: prospects for an end to deforestation.  
316 *Annu. Rev. Environ. Resour.* **36**, 343–371 (2011).
- 317 55. Brashares, J. S., Golden, C. D., Weinbaum, K. Z., Barrett, C. B. & Okello, G. V.  
318 Economic and geographic drivers of wildlife consumption in rural Africa. *Proc. Natl.*  
319 *Acad. Sci.* **108**, 13931–13936 (2011).
- 320 56. Allen, J. C. & Barnes, D. F. The causes of deforestation in developing countries. *Ann.*  
321 *Assoc. Am. Geogr.* **75**, 163–184 (1985).
- 322 57. World Bank. The World Bank Databank. (2013). Available at: [databank.worldbank.org](http://databank.worldbank.org).
- 323 58. FAO and JRC. *Global forest land-use change 1990-2005. FAO Forestry Paper 169, Food*  
324 *and Agriculture Organization of the United Nations*. (2012).
- 325 59. FAO. *Global Forest Resources Assessment 2010: Main Report: FAO Forestry Paper 163,*  
326 *Food and Agriculture Organization of the United Nations*. (2010).
- 327 60. Hansen, M. C. *et al.* High-Resolution Global Maps of 21st-Century Forest Cover Change.  
328 *Sci.* **342**, 850–853 (2013).
- 329 61. Redding, D. W. & Mooers, A. Ø. Incorporating Evolutionary Measures into Conservation  
330 Prioritization. *Conserv. Biol.* **20**, 1670–1678 (2006).
- 331 62. Male, T. D. & Bean, M. J. Measuring progress in US endangered species conservation.  
332 *Ecol. Lett.* **8**, 986–992 (2005).
- 333 63. Clemens, M. A., Radelet, S., Bhavnani, R. R. & Bazzi, S. Counting chickens when they  
334 hatch: timing and the effects of aid on growth. *Econ. J.* **122**, 590–617 (2011).
- 335 64. Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A. & Smith, G. M. *Mixed effects*  
336 *modelling and extensions in ecology with R*. (Springer, 2009).
- 337 65. Balmford, A. Extinction filters and current resilience: the significance of past selection  
338 pressures for conservation biology. *TREE* **11**, 193–196 (1996).
- 339 66. Cragg, J. Some statistical models for limited dependent variable with application to the  
340 demand for durable goods. *Econometrica* **39**, 829–844 (1971).



- 341 67. Joergensen, B. Exponential dispersion models. *J. R. Stat. Soc. B* **49**, 127–162 (1987).
- 342 68. Joergensen, B. *Theory of Dispersion Models*. (Chapman and Hall, 1997).
- 343 69. Tweedie, M. An index which distinguishes between some important exponential families.
- 344 in *Statistics: Applications and New Directions. Proceedings of the Indian Statistical*
- 345 *Institute Golden Jubilee International Conference 579–604* (Indian Statistical Institute,
- 346 1984).
- 347 70. Zhang, W. cplm: Compound Poisson Linear models version 0.7-2. Available at the
- 348 Comprehensive R Archive Network (CRAN). (2014).
- 349 71. Burnham, K. P. & Anderson, D. R. *Model selection and multimodel inference: a practical*
- 350 *information-theoretic approach*. (Springer-Verlag, 2002).
- 351 72. Cohen, J. *Statistical power analysis for the behavioural sciences*. (Laurence Erlbaum,
- 352 1988).
- 353 73. Ferraro, P. & Hanauer, M. Advances in measuring the environmental and social impacts
- 354 of environmental programs. *Annu. Rev. Environ. Resour.* 495–517 (2014).
- 355 74. Imai, K. & Ratkovic, M. Covariate Balancing Propensity Score. *J. R. Stat. Soc. B* **76**, 243–
- 356 263 (2012).
- 357 75. Fong, C., Ratkovic, M. & Imai, K. CBPS: R package for covariate balancing propensity
- 358 score. available at the Comprehensive R Archive Network (CRAN). (2015).
- 359 76. Pandit, R. & Laband, D. N. Spatial autocorrelation in country-level models of species
- 360 imperilment. *Ecol. Econ.* **60**, 526–532 (2006).
- 361 77. Amin, A. & Choumert, J. Development and biodiversity conservation in Sub-Saharan
- 362 Africa: A spatial analysis. *Econom. Bull.* **35**, 729–744 (2015).
- 363 78. Drukker, D. M., Prucha, I. R. & Raciborski, R. On two-step estimation of a spatial
- 364 autoregressive model with autoregressive disturbances and endogenous regressors.
- 365 *Econom. Rev.* **32**, 686–733 (2013).
- 366 79. R Core Team. R: A language and environment for statistical computing. (2013).
- 367 80. Fox, J. *An R and S-Plus companion to applied regression*. (Sage Publications, 1997).
- 368 81. IUCN. *The IUCN Red List of threatened species*. (2009).
- 369 82. Schipper, J. *et al.* The status of the world’s land and marine mammals: diversity, threat,
- 370 and knowledge. *Science* **322**, 225–30 (2008).
- 371 83. Birdlife International. Species distribution data download. (2014).
- 372 84. Sandvik, B. World Borders Dataset. (2009). Available at:
- 373 [http://thematicmapping.org/downloads/world\\_borders.php](http://thematicmapping.org/downloads/world_borders.php).
- 374 85. Pebesma, E. J. & Bivand, R. S. Classes and methods for spatial data in R. *R News* **5**,
- 375 (2005).
- 376 86. Bivand, R. . & Kewin-Koh, N. Maptools: Tools for reading and handling spatial objects. *R*
- 377 *package version 0.8-27* (2013). Available at: <http://cran.r-project.org/pacakge=maptools>.
- 378 87. Balmford, A., Gaston, K. J., Blyth, S., James, A. & Kapos, V. Global variation in
- 379 terrestrial conservation costs, conservation benefits, and unmet conservation needs. *PNAS*
- 380 **100**, 1046–1050 (2003).

381  
382 **Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

383  
384  
385  
386  
387  
388

389 **FIGURE LEGENDS**

390  
391

392 **Figure 1 | Global biodiversity declines and conservation spending impacts.** Colours show  
393 percentage of all global declines (total BDS) associated with each country: dark red = >10%  
394 (Indonesia only, 21%); dark, mid and light orange = 5–10, 2.5–5 and 1–2.5% respectively;  
395 yellow = 0–1%; grey indicates BDS = 0; blue indicates a net improvement in national  
396 biodiversity status. Pies show the predicted reduction in decline (in black) if spending had  
397 been i\$5million higher (for selected countries); pie size represents  $\sqrt{\text{BDS}}$ . Inset shows  
398 predicted vs. observed BDS (ln-transformed) for the continuous model (see also Extended  
399 Data Figure 4). Country outlines supplied by `esri_dm`  
400 [https://www.arcgis.com/home/item.html?id=d86e32ea12a64727b9e94d6f820123a2#ov](https://www.arcgis.com/home/item.html?id=d86e32ea12a64727b9e94d6f820123a2#overview)  
401 [erview](https://www.arcgis.com/home/item.html?id=d86e32ea12a64727b9e94d6f820123a2#overview)

402  
403  
404

405 **Figure 2 | The country-scale rate of biodiversity decline (BDS) depends on conservation**  
406 **spending levels.** The continuous part of the model is shown (which focuses on high-decline  
407 countries, n=50 independent countries) and both variables are corrected for all other predictors in  
408 a residual-residual plot (Pearson’s  $r = -0.69$ ). See Table 1 for spending impact in the binomial  
409 model part.

410  
411

412 **Figure 3 | Conditional impacts of human pressures on biodiversity.** (a) Impact of GDP  
413 growth on BDS depends on the existing level of GDP/capita. Red = slow GDP growth ( $10^{\text{ile}}$ ),  
414 blue = fast growth ( $90^{\text{ile}}$ ), “low” GDP/capita =  $10^{\text{ile}}$ , “median” =  $50^{\text{ile}}$  (effects are still significant  
415 at  $>50^{\text{ile}}$ ). (b) Impact of agricultural expansion on BDS depends on the existing % of land  
416 converted to agriculture: colours as in (a), “low” agricultural expansion =  $10^{\text{ile}}$ , “median” =  $50^{\text{ile}}$   
417 (effects are still non-significant at  $<50^{\text{ile}}$ ). Error bars show conditional 95% confidence intervals  
418 from the continuous model-part. N=50 independent countries. Centre is the median.

419  
420

421 **Table 1 | Best-fit models predicting biodiversity decline.** Note that for all terms that  
 422 interact, the interaction plots provided must be used to interpret the reported  
 423 standardised coefficients correctly (Figure 3 and Extended Data Figures 2-3). “Agric.  
 424 land” = mean percentage of agricultural land; t-1 = 1994-2000, t-2 = 1988-1994; GDP  
 425 = Gross domestic product per capita PPP; population = rural population density;  
 426 governance improvement = change in the government effectiveness score. N=50  
 427 independent countries and index parameter = 1.01 in the continuous part, n=109  
 428 countries in binomial part with a 42:67 ratio of ones to zeroes.  
 429

Predictor variable	Continuous model part (BDS)	Binomial model part (BDSb)
Conservation spending	-0.251	-4.800
Agricultural growth	-0.012	-3.065
GDP growth	0.035	-0.152
Population growth	NA	-2.738
Declines in period t-1	0.024	NA
Declines in period t-2	0.048	NA
Threatened species richness	0.155	5.421
Country area	NA	8.754
GDP	0.037	-5.426
% agric. land	0.049	-1.226
GDP growth x GDP	0.031	NA
Spending x GDP	NA	5.026
Spending x threatened species richness	-0.247	NA
Population growth x GDP growth	NA	1.044
Agric. growth x % agric. land	-0.045	-10.143
Spending x % agric. land	0.065	NA
Agric growth x governance improvement	NA	-9.603

430

431

433 **Materials and Methods**434 Country-scale biodiversity decline scores

435 To quantify biodiversity decline, we used equally-weighted genuine changes in  
 436 the IUCN Red List status of all global bird and mammal species up to the last Global  
 437 Mammal Assessment in 2008 (i.e. changes in extinction risk between 1996-2008 for  
 438 mammals and 2000-2008 for birds, there being no 1996 global bird assessment; the  
 439 term “genuine” excludes any Red List changes not related to changing extinction  
 440 risk, in particular those due simply to taxonomic changes)<sup>3,31,32</sup>. Our approach is  
 441 therefore similar to planet-scale Red List Indices (RLI) of global biodiversity change  
 442 adopted by governments to measure performance against CBD and SDG targets<sup>3,31-33</sup>,  
 443 but adjusted to allow global declines to be portioned out among signatory countries  
 444 while preserving the original magnitude of declines. We focused on birds and  
 445 mammals because these received the vast majority of conservation investment and  
 446 supply robust, directly-observed data on changes in Red List status<sup>3,34</sup>; thus, we  
 447 excluded the other possible taxon (amphibians) because they received almost no  
 448 conservation investment during the study period<sup>3</sup>, only have modelled (rather than  
 449 directly observed) declines available for 1980-2004<sup>3,35</sup> (whereas robust spending data  
 450 are only available from 1992 onwards<sup>8</sup>), and are also highly data deficient and  
 451 “enigmatic” in terms of their declines<sup>3,35</sup>.

452 To convert species-based Red List changes into country-level indices of  
 453 biodiversity change, we divided up each species change as “decline fractions”, based  
 454 on the percentage of the species range  $p_{ij}$  held by each country<sup>8,26</sup>. However, decline  
 455 fractions are estimates for the underlying responsibility fraction  $R_{ij}$  = the proportion of  
 456 the status change for species  $i$  attributable to country  $j$  (see Additional Method Details  
 457 at end). For greater accuracy, we therefore corrected these range-based fractions in  
 458 two ways. First, the losses underlying a species decline are not homogeneously  
 459 distributed in space but instead, are frequently concentrated in some part of the range  
 460 where human pressures have suddenly increased<sup>36</sup>. Both empirically and at random,  
 461 those concentrations of pressure-driven loss are unlikely to lie at the range periphery  
 462 (Additional Method Details and<sup>36</sup>). However, a raw range-based algorithm assumes  
 463 spatially homogeneous losses right up to the range periphery, and so will often assign  
 464 an erroneous and trivial responsibility fraction to any country holding a small range-  
 465 edge ( $p_{ij}$ ) of a species found almost entirely in a neighbouring country. Formally,  $R_{ij}$   
 466 for small  $p_{ij}$  is often but not always likely to be zero. These small  $p_{ij}$  values were also  
 467 extremely numerous, generating very high noise-to-signal ratios in analysis. To  
 468 address these problems of extreme signal loss and bias when an unknown proportion  
 469 of small  $p_{ij}$  were incorrect overestimates of zero  $R_{ij}$ , we used Signal Detection  
 470 Theory<sup>37</sup> and the mathematics of the Red List categories to estimate a range of  
 471 theoretically optimal thresholds  $T$  such that  $R_{ij}$  is set to zero if  $p_{ij} < T$ , and then carried  
 472 out our analyses using three possible thresholds within this range, to account for  
 473 uncertainty (see Additional Method Details). The main text shows results for  $T=0.17$   
 474 (being the approximate optimal trade-off between noise reduction and sample size,  
 475 Additional Method Details) and Supplementary Results and Extended Data Table 4  
 476 shows sensitivity tests with alternative thresholds (including the finding that  
 477 explanatory power at  $T=0.17$  is considerably stronger than occurs with the other  
 478 thresholds).

479 Second, we analysed the Red List reports for each individual bird and mammal  
 480 species and altered the range-based fractions wherever a report suggested a different

481 distribution of responsibilities across countries (Supplementary Table 4). We then  
482 calculated the Biodiversity Decline Score (BDS) for each country by summing all  
483 decline fractions for birds and mammals, treating the rare status improvements as  
484 negative fractions<sup>8,26</sup>. Supplementary Table 1 contains the final BDS scores per  
485 country.

#### 487 Predictors of country-scale biodiversity decline scores (BDS)

488 Conservation models with policy relevance need to have general applicability,  
489 including being able to accurately forecast outcomes when presented with situations  
490 that are different from the original dataset on which they were parameterised. To  
491 achieve this, it is highly advisable to use broad, general variables because more  
492 specific ones often have very poor forecasting performance when used beyond the  
493 original data<sup>37</sup>. We therefore selected a relatively small set of simple, generalised and  
494 publicly available explanatory variables to represent national-level socioeconomic  
495 pressures, noting that conservation spending also captures overall conservation effort  
496 in a broad, quantifiable and publicly-reported way.

497 For conservation spending, we took data on average annual conservation  
498 investment levels from a recently-published collation<sup>8</sup>, adding new data for countries  
499 that had been data-deficient in the original published study e.g. Turkey<sup>38</sup>. Finance data  
500 were collated at 2005 constant U.S. dollar values (consistent with<sup>8</sup>) but for analysis,  
501 were converted to “international dollars” (abbreviated as i\$ in the main text) at local  
502 purchasing power parity values, where purchasing power parity accounts for  
503 differences in the purchasing power of U.S. dollars (when exchanged) in each  
504 country<sup>39</sup>. Two types of conservation investment data were available: (a) “strict-  
505 sense” funding with direct links to biodiversity conservation, and (b) “mixed funding”  
506 mainly targeted at social and development goals but with potential indirect, long-term,  
507 and often unclear impacts on biodiversity (e.g. school-building or agricultural  
508 assistance in forest communities)<sup>28,40</sup>. *A priori*, we hypothesized that strict-sense  
509 biodiversity funding was likely to be the better predictor of rates of decline, whereas  
510 “mixed” development funding (which involves much larger sums than strict-sense  
511 funding) was likely to obscure any effect. “Strict-sense” funding also produced lower  
512 AICc scores in exploratory modelling, and so we used it in our final analysis.

513 Good governance is also hypothesized to positively affect biodiversity, both  
514 directly (e.g. through reducing conflict) and indirectly (e.g. through making  
515 conservation investment more efficient)<sup>16,41–43</sup>. Governance has been measured using  
516 multiple indicators<sup>44</sup>, so we modelled the impact of change in the six indicators  
517 published in the World Governance Indicators dataset<sup>44</sup>: government effectiveness,  
518 political stability and conflict, rule of law, corruption, regulatory quality (largely a  
519 measure of openness to business activity) and “voice” (a measure of the democratic  
520 accountability of governments). All the governance indicators are very tightly  
521 correlated with each other ( $r > 0.9$  for all pairwise combinations) and so to avoid  
522 collinearity, we tested each one individually. Government effectiveness gave the best  
523 fit in exploratory analysis (as in<sup>8</sup>) and is reported in the results as “governance”.

524 For the country-level pressures aspect of our PACI model, we followed previous  
525 authors in using national rates of human population growth, economic growth and  
526 agricultural expansion<sup>13–15,17,45–53</sup>. Such country-level aggregators likely capture the  
527 overall impact of multiple smaller-scale drivers (with agriculture being the main  
528 pressure driving threat<sup>24</sup>). For example, forest clearance for food production or  
529 commodities would generally cause changes in both area of agricultural land and  
530 economic output, and GDP levels have been associated with both hunting pressure

531 and deforestation trends<sup>54-56</sup>. For economic growth, we used change in GDP/capita  
532 PPP (purchasing power parity). For agricultural growth, we used change in the  
533 percentage of land converted to agriculture; and for population growth, we used  
534 change in human population density (using total and rural population density as  
535 alternatives). Data on GDP, agricultural land and human populations were taken data  
536 from World Bank statistical tables<sup>57</sup>. We also tested the direct impact of forest loss,  
537 estimated per country for 1990-2000 using FAO statistics<sup>58,59</sup> (although we  
538 acknowledge the limitations of this historical dataset<sup>60</sup>).

539 The number of declining species in a country (and hence its BDS) is likely to be  
540 strongly influenced by the total number of species present and/or country area, plus  
541 the starting-condition levels of risk and decline. Following previous studies (e.g.<sup>61</sup>),  
542 we calculated total threatened species richness in the same way as we calculated total  
543 species decline (BDS), i.e. we summed all species fractions in each country,  
544 weighting them by the level of extinction risk as an index of threat. We compiled  
545 country area from<sup>8</sup>. However, in exploratory analysis, we found that the inclusion of  
546 area in any continuous-part model consistently led to a worse fit (delta AICc >6.5),  
547 likely because species richness absorbed most of the variance explained by area in  
548 this (n=50 countries) sample. In contrast, binomial-part models (n=109 countries)  
549 detected separate area and species richness effects (without collinearity; Extended  
550 Data Tables 2-3). Thus, we included the area term in binomial models, but excluded  
551 it from our final set of continuous-part models. We note, however, that parameter  
552 estimates with and without area were extremely similar.

#### 553 554 Lags between predictors and responses

555 Conservation investment/action takes at least 5 years, and often over a decade, to  
556 have an impact on biodiversity<sup>29,62</sup>, especially for taxa such as birds and mammals.  
557 For mammals, the two global Red List assessments from which status changes can be  
558 calculated were in 1996 and 2008<sup>3</sup>. We therefore assumed that changes detected in the  
559 2008 assessment may have been driven by conservation finance allocations occurring  
560 as recently as five years earlier (i.e. 2003) but in all likelihood, could also be  
561 influenced by spending from a decade or more earlier (the early 1990s). Similarly,  
562 changes occurring after 1996 (i.e. starting in 1997) could have been influenced by  
563 spending allocations as early as 1992 (also the year in which global conservation  
564 spending began in earnest with the Rio Earth Summit<sup>28</sup>) but also by allocations up to  
565 the early 2000s. Following this logic, we used predictor variables for 1992-2003  
566 (annualised values) to model changes in the response value for 1996-2008, using the  
567 same lag for the four different socioeconomic growth variables to avoid the analysis  
568 becoming intractable. We tested an alternative predictor period of 1992-2000 but  
569 preferred 1992-2003 based on lower AICc values.

570 Technically, therefore, our response variable is a lagged variable<sup>63</sup> taking the  
571 form  $Y_t - Y_{t-n}$  and our socioeconomic change variables are similarly lagged. We  
572 acknowledge that predicting change occurring in a time block using variables from an  
573 earlier time block is necessarily approximate, but year-by-year species changes were  
574 not available. Nevertheless, country-level patterns of change in predictor variables  
575 were strongly correlated across different time periods (e.g. when comparing mean  
576 annual values for 1992-2000 and 1992-2003, the correlations for population growth,  
577 population size, GDP growth and GDP respectively are 0.91, 0.999, 0.89, and 0.999).  
578 These strong correlations imply that the precise choice of year/period seems unlikely  
579 to have an important effect on the results.

580 The rate of decline over a fixed period is also likely to be influenced by the  
581 “inertia” from declines in the years immediately preceding that period. To explore  
582 this, we calculated avian BDSs for the two IUCN assessment periods preceding our  
583 study period (1988-1994 and 1994-2000) and added both measures to our candidate  
584 regression models. No earlier-period BDS was available for mammals; however,  
585 mammal and bird BDS are highly correlated in the study period (Pearson’s  $r = 0.998$ ),  
586 so we assumed earlier-period bird BDSs to be reasonable proxy of combined (bird +  
587 mammal) earlier-period BDSs.

588

### 589 Statistical analysis

590 All predictor variables were z-standardized to put effect sizes on a common  
591 scale<sup>64</sup>. We excluded any countries for which complete, robust data were lacking  
592 (see<sup>8</sup>), including where reported finance commitments cannot be safely regarded as  
593 strict-sense biodiversity spending. We also excluded countries that had multiple  
594 overseas territories but conservation spending was not disaggregated across those  
595 territories, despite strongly different values for the socioeconomic predictors and rates  
596 of decline across the territories. In particular, the USA, France and the UK were  
597 excluded from regression models under this rubric (and we therefore recommend  
598 greater geo-referenced finance reporting). See Supplementary Table 1 for all  
599 exclusions. The Solomon Islands and New Zealand represented potentially influential  
600 leverage points, so we tested models both with and without these countries. We found  
601 that inclusion of the Solomon Islands had a large impact on binomial outcomes  
602 (causing governance growth to be dropped from the best-fit binomial-part model,  
603 likely due to the extreme value of governance growth for the Solomons), so we  
604 excluded this country from all binomial models. The impact of including the Solomon  
605 Islands was smaller in the continuous part (an identical best-fit model with similar  
606 coefficients was selected whether the country was included or excluded) but for  
607 completeness, we consistently tested all continuous model variations both with and  
608 without the Solomons. Inclusion of New Zealand had a major impact on binomial-part  
609 outcomes, altering most coefficients by ~20% and some by >100%, and also greatly  
610 worsened fit in the continuous part, so it was excluded overall. The leverage  
611 associated with including New Zealand may be due to this country having a negative  
612 value for agricultural growth.

613 We then built candidate PACI models to predict BDS, each testing hypotheses  
614 about how conservation investment and various human pressures might impact on  
615 biodiversity (see Supplementary Table 2 for full list). We included several  
616 interactions to test whether socioeconomic context altered the impact of  
617 socioeconomic change. For example, we hypothesized that in countries that have  
618 already converted much of their land base to agriculture, additional expansion of  
619 farmland may either have a reduced marginal effect on biodiversity due to an  
620 extinction filter<sup>65</sup>, or a greater impact as the last vestiges of habitat disappear  
621 (Supplementary Discussion). Thus, we further calculated mean annual values of GDP,  
622 population, governance and % agricultural land for 1992-2003 and added these to our  
623 interaction model specifications. Extended Data Table 1 and Supplementary Table 2  
624 show all interactions tested.

625 The BDS data were non-integer covering both positive and negative values, but  
626 had a relatively dense cloud of values at zero. Although a more limited number of  
627 zeroes does not violate regression assumptions, such a long tail of zeroes can generate  
628 extreme bias<sup>64</sup>. We therefore used the recommended approach of a two-part model<sup>27,66</sup>  
629 that creates (a) a “continuous” part ( $n=50$  countries) comprising all countries with a

630 non-zero BDS plus informative zeroes; (b) a “binomial” part (n=109) that included all  
631 countries with data (and so all zeroes), but converted BDS to the binary response  
632 BDSb (where  $BDSb = 1$  if  $BDS > 0$  and 0 otherwise). For the continuous part  
633 specifically, we sought to optimise the trade-off between information content and bias  
634 by including as many zeroes as possible, in order of their likely informativeness,  
635 without causing clear patterns in regression diagnostic plots (thus extending the  
636 principle of the hurdle models developed for non-negative integer data<sup>64</sup> to two-part  
637 analyses). A country that has many species but has experienced no declines, such as  
638 Costa Rica, suggests an important underlying process captured by zero BDS (= higher  
639 informativeness of zero decline). Conversely, when a country is species-poor, there is  
640 a strong random expectation that over a 13-year period, no species will be observed  
641 changing its Red List status (= lower informativeness of the zero). We therefore  
642 defined  $\Psi$  as country-level species richness (derived from our prior geographic  
643 analysis) and then, for various possible values of this parameter, heuristically tested  
644 the degree of regression bias arising when we excluded all cases of  $\{BDS=0$  and  
645 species richness  $< \Psi\}$ . We found a tradeoff whereby setting  $\Psi$  at 40 or more left  
646 minimal patterns in residual plots but reduced sample size and statistical power,  
647 whereas  $\Psi$  values below 20 started to generate strong patterns in plots of residuals  
648 against fitted values. We therefore chose a value of  $\Psi = 25$  (see Supplementary  
649 Results and “Sensitivity Testing” (below) for sensitivity testing on this parameter).

650 For the continuous part, BDS retained a right skew even after log-transformation  
651 (Extended Data Figure 5) and there was also heteroscedasticity in the errors, so we  
652 tested Generalized Linear Models (GLMs) with the gamma-like Tweedie error  
653 distribution, which uses maximum likelihood to simultaneously model heteroscedastic  
654 variance as a function of the mean<sup>67-69</sup> (cplm R package<sup>70</sup>). We carried out an (X+10)  
655 transformation on BDS to avoid violating gamma assumptions (where the value of 10  
656 was chosen to give flexibility for modelling with future scenarios where more species  
657 recoveries may occur, and where BDS may therefore become more negative).  
658 Tweedie model selection often uses the Gini index for model selection<sup>70</sup>. However,  
659 the ratio of sample size to the number of parameters is relatively small in the Tweedie  
660 analyses, potentially indicating low power to distinguish among models and a risk of  
661 overfitting. Thus, we initially compared model fit using the Gini index, but then re-ran  
662 model selection using AICc, a technique which penalizes overfitting and is  
663 asymptotically similar to leave-one-out cross validation<sup>71</sup>, and regarded Gini-selected  
664 models as overfitted if they contained terms that both were excluded in AICc  
665 selection and had  $p > 0.1$ . Gini and AICc approaches gave identical model selection  
666 results in the main text; in the sensitivity tests for  $T=0.10$  and  $T=0.25$ , however (see  
667 Sensitivity Testing, below), we preferred AICc approaches. We also carried out a  
668 power analysis<sup>72</sup>, which revealed that our best-fitting models had a power of  $>0.99$ ,  
669 and thus that our sample size was adequate to detect effects among the relatively large  
670 number of parameters.

671 In the binomial part, exploratory GAMs again suggested that linear modelling  
672 was appropriate, and so we used GLMs with binomial errors, fitting an additional  
673 dispersion parameter to account for strong underdispersion<sup>64</sup>. Models containing this  
674 extra parameter do not generate AIC values, so we carried out non-automated  
675 binomial model selection, using stepwise backward and forward regression with  
676 likelihood ratio tests<sup>64</sup>. Explanatory power was measured in the continuous part using  
677 McFadden’s  $R^2$  (known to be conservative), and in the binomial part using the  
678 percentage of times that the model correctly predicted BDSb (taking  $p(BDSb=1)$   
679  $< 50\%$  as a predicted 0, and  $p(BDSb=1) > 50\%$  as a predicted 1).



680

681 Cross validation to test for forecasting accuracy on unseen data

682 To test the model's forecasting accuracy, as would be needed for policy  
683 usefulness, we carried out ten-fold cross-validation, a procedure that repeatedly sets  
684 aside part of the data (as a "fold" of BDS values the model has never seen),  
685 parameterises the model on the remaining subset of data, then tests how well it  
686 forecasts the unseen BDS values<sup>37</sup>. For the continuous model part, we measured  
687 forecasting accuracy by calculating McFadden's  $R^2$  for the model fit to the unknown  
688 (hold-out) BDS in each of the ten folds. Ideally, the slope of forecast versus known  
689 values should also be close to 1.0 and to test for this, we regressed the complete set of  
690 forecast values (across the ten folds) against the complete set of known values in the  
691 cross-validation, using a Generalized Least Squares regression model with a constant  
692 power function fitted to describe the heteroscedasticity in the residuals. We also  
693 calculated the median absolute deviation, although this is less informative in data with  
694 a large spread of values (note also that percentage deviations, rather than absolute  
695 deviations, are not appropriate metrics for low-volume data containing several zeroes  
696 such as BDS<sup>37</sup>). For the binomial model part, we tested mean forecasting accuracy  
697 against unknown data using % correct predictions, as we had done in testing binomial  
698 explanatory power.

699

700 Covariate balancing and spatial considerations

701 An important issue with impact studies is "selection bias", where the likelihood  
702 of receiving the intervention of interest is non-random<sup>25,73</sup>. The amount of  
703 conservation investment a country receives is indeed known to be influenced by non-  
704 random factors including Red List status itself<sup>8</sup>, potentially creating endogeneity  
705 problems<sup>25,73</sup> and in particular, a potential problem of reverse causality whereby  
706 decline drives changes in conservation spending rather than vice versa.

707 Our use of a time lag between predictors and responses was designed to reduce  
708 the issue of reverse causality in the analysis. We also note that since greater decline  
709 has been shown to cause greater investment<sup>8,28</sup>, a simple reverse-causality hypothesis  
710 would imply a positive correlation between spending and decline, whereas we  
711 observed a negative correlation (greater investment was associated with less  
712 subsequent decline). To correct for selection bias and associated endogeneity  
713 problems more generally, we used covariate balancing propensity scores<sup>25</sup> for  
714 continuous treatment variables<sup>74</sup> (in the R package CBPS<sup>75</sup>), which minimises the  
715 association (the Pearson correlation) between covariates and the treatment<sup>74,75</sup>.  
716 Previous studies have explained a high proportion of the variance in conservation  
717 finance allocation using country area, cost (the National Price Level), government  
718 effectiveness, political stability, GDPPPP, the percentage of land that is a protected  
719 area, and the sums of threatened bird and mammal species weighted by their level of  
720 extinction risk<sup>8,28</sup>. We carried out covariate balancing using data on these variables  
721 (taken from<sup>8</sup>) plus data on forest loss between 1990 and 2005 (taken from the FAO  
722 data<sup>58,59</sup>) and data on 1992-2003 growth in GDP per capita PPP (taken from World  
723 Bank data<sup>57</sup>). Extended Data Figure 6 shows the Pearson correlations between the  
724 treatment and the covariates before and after the covariate balancing propensity score  
725 correction.

726 Analysing species declines at the country level could potentially generate spatial  
727 structure in model residuals, violating regression assumptions<sup>50,64,76,77</sup>. We tested for  
728 this effect by fitting four possible structures to the most complete GLM model using  
729 REML (restricted maximum likelihood estimate) and comparing their predictive

730 power using AICc. The structures tested were: (i) a fixed effect for Region (see<sup>8</sup> and  
731 Supplementary Tables 1-2 for regions and regional intercept differences); (ii) a  
732 GLMM with a SAC (Generalized Additive Mixed model with spatial autocorrelative  
733 structure), where five possible structural models describing the spatial autocorrelative  
734 structure between country centroid coordinates were tested – linear, spherical,  
735 Gaussian, ratio and exponential<sup>64</sup>; (iii) a GLMM with an SAC as in (ii) plus a fixed  
736 effect for Region; (iv) a GLMM with an SAC plus a random intercept for Region. The  
737 best-fitting structure was (i) and we used this in subsequent modelling. Using Region  
738 as a fixed effect also follows logically from theory, since regional differences are a  
739 potentially important component of decline<sup>46</sup>. Binomial models including spatial  
740 autocorrelative structures did not converge and regional effects were non-significant,  
741 so we tested for possible spatial effects by plotting residuals from the best-fit binomial  
742 model against both latitude and longitude, and also by exploring the effect of  
743 including the latitude and longitude coordinates of the country centroids in the model  
744 specification. There was no support for models including latitude and longitude and  
745 no visual relationship in the plots against residuals.

746 Decline drivers in one country may have impacts on biodiversity in neighbouring  
747 countries and statistical “spatial lags” have been used to model such possible  
748 effects<sup>50,77</sup>. However, statistical techniques to model a mixture of spatial error and  
749 spatial lag in the dependent and independent variables have only recently been  
750 developed for OLS regression<sup>78</sup> and to our knowledge, no robust methodology exists  
751 for non-linear generalized models with heteroscedastic Tweedie error structures. We  
752 therefore restricted ourselves to testing and correcting for spatial error structures.  
753 However, by dividing responsibility for declines proportionally among countries, we  
754 have likely removed much of the artefactual spatial lag that arises when neighbouring  
755 countries are given equal responsibility for any declining species that they share.

756 All statistical analysis was carried out in the R statistical software environment<sup>79</sup>.  
757 We checked for violations of model assumptions using diagnostic plots of residuals  
758 against fitted values and against all candidate predictors variables<sup>64</sup>. When removing a  
759 variable in model selection, we also plotted the residuals of each reduced model  
760 against the newly-removed variable, checking for any pattern that the statistical tests  
761 may have missed. Collinearity was checked for using VIF scores (Extended Data  
762 Table 3).

#### 763 764 Predicting the impact of spending and pressure changes

765 To predict the impact that an extra i\$1m or i\$5m dollars annually of conservation  
766 spending would have had in each country, we added those amounts to known  
767 financing levels for each country and used the model to re-predict the outcomes. To  
768 predict the effect of changing human pressures on those outcomes, we followed the  
769 same protocol but also replaced the 1992–2003 levels of socioeconomic growth (i.e.  
770 change in pressures) with 2001–2012 levels. To estimate the decline that we may have  
771 avoided as a result of 1992–2003 spending, we used the fact that prior to the 1992  
772 Earth Summit, biodiversity spending for which we have data was flat and often zero  
773 (noting that data becomes sparser prior to the 1990s, and sparser still as one goes back  
774 further in time). We therefore estimated mean annual spending for 1985–1990, then  
775 re-predicted outcomes as if post-1990 annual budgets had only increased in line with  
776 inflation (i.e. no real increase). Although reduced data quality and imputation for the  
777 1985–1990 spending makes these estimates approximate, the median change in BDS  
778 was robust to several different spending estimates, and so the global figure for

779 avoided decline (29%) is likely to be a reasonable approximation, although we  
780 acknowledge that the true figure may be higher or lower.

781

### 782 Sensitivity Testing

783 We further tested the sensitivity of our original PACI model to various  
784 assumptions. To test for sensitivity to the threshold  $T$  (which was set at 0.17 in the  
785 main text, see Additional Method Details, below), we examined the model outcomes  
786 using  $T = 0.10$ . and  $T = 0.25$ . To test for sensitivity to the  $\Psi$  parameter, we repeated  
787 the analysis with multiple variations around the parameter value used in the main  
788 analysis, finding no qualitative differences in the results. To test for the effect of the  
789 influential outliers (Solomon Islands and New Zealand), we ran model selection both  
790 with and without the outliers. To examine whether our results were sensitive to the  
791 variables used to calculate the propensity scores (the correction for non-random  
792 assignment of spending amounts across countries, see “covariate balancing and spatial  
793 considerations” above), we tested the impact of removing various individual variables  
794 or combinations of variables from the list used to calculate the propensity weights for  
795 the regression model.

796 A further concern was that our model fits might be driven (biased) by a country  
797 or countries with high BDS, since the BDS distribution is skewed (Extended Data  
798 Figure 5). Our tenfold cross-validation test already showed that the omission of  
799 various groups of countries had no substantive impact on results but as a further  
800 check, we carried out a jack-knife leave-one-out test to see how the omission of each  
801 individual country affected parameter estimates. When interactions between  
802 continuous terms are present, parameter estimates are conditional, i.e. they are  
803 different for each country and indeed affect each other. An appropriate measure of  
804 parameter change is therefore the average percentage change in the values of the  
805 conditional expectations across all countries. For example, if a country  $C$  (such as  
806 Indonesia) was strongly biasing the model results, then when we re-run the model  
807 without  $C$ , we should see a substantial change in the average conditional expectation  
808 of BDS across the remaining countries, indicating a strong shift across the conditional  
809 parameter estimates for the interaction model. With heteroscedastic errors, the median  
810 percentage may also be more informative than the mean, so we considered both.

811 Even with these tests, there remained the possibility of “joint influence” in the  
812 continuous model part<sup>80</sup> where the highest-value BDS countries were driving the  
813 model as a group (for example, the BDS values for the top three countries of  
814 Indonesia, Australia and China are very large, being 272%, 69% and 24% larger than  
815 the fourth-highest BDS value, and so may combine to exert joint leverage on the  
816 model parameters). To test for this, we plotted fitted against observed values for both  
817 the full dataset and the top-three-removed dataset. For completeness, we also  
818 examined changes in the individual conditional coefficients when the top three BDS  
819 countries were omitted.

820 In impact assessments addressing the impact of a single variable, a further  
821 concern is “missing variable bias”, where there may be a confounding variable closely  
822 correlated with both the studied impact variable and the outcome variable<sup>87</sup>. In other  
823 words, the observed impact of conservation spending may simply be an artefact of  
824 spending being collinear with an unknown variable that is actually driving the  
825 outcome. When only one explanatory variable is being studied for its impact, hidden  
826 variable bias can be investigated by testing whether the main variable impact is still  
827 observed after an artificially created, collinear dummy variable has been added to the  
828 analysis<sup>25</sup>. In multiple regression analyses, this is largely infeasible because it would

829 also be necessary to artificially generate correlations between the dummy and all the  
830 other (interacting) variables in the regression formula. Nevertheless, we attempted to  
831 take the spirit of the missing variable test by looking for an empirical variable that  
832 was closely correlated with our spending variable (and therefore had a natural co-  
833 correlation with all other variables in the regression formula), then adding it into the  
834 regression and testing whether the spending impact disappeared. Using the same  
835 scaling standardization as in the main analysis, we found that mean total population  
836 size had a correlation (Pearson's r) of 0.45 with spending and mean GDPPPP (i.e. raw  
837 GDP rather than the GDP per capita used in the main analysis) had a correlation of  
838 0.54 with spending. We therefore tested the impact of adding both variables in turn to  
839 our regression formulae (in the second instance, removing GDP per capita and  
840 replacing it with raw GDP, on account of a strong correlation between the two).

841 Finally, we tested the possible impact of inaccuracy in national conservation  
842 spending data, following the sensitivity tests used in<sup>8</sup>: in summary, we varied the  
843 spending data for each country by iteratively drawing new spending values for each  
844 country from a normal distribution centred on the original value and with a standard  
845 deviation set to 25% of the original value, and then repeating the regression analysis.  
846 Owing to extremely slow processing times for our complex models, we carried out  
847 100 such permutations.

848 Detailed results of all these sensitivity tests are shown in the Supplementary  
849 Results, but none affected our conclusions substantively.

850

851

#### 852 Additional Method Details: Mathematical calculation of BDS

853 Although change in Red List status is a standard measure of biodiversity change  
854 used in the CBD and SDG frameworks<sup>3,31,32</sup>, it applies to species, whereas we wished  
855 to measure change at the level of the sovereign countries that, as signatories to these  
856 agreements, have the principal political responsibility for biodiversity policy and  
857 targets. We therefore created an algorithm to convert species-level change to country-  
858 level change. Mathematically, we define  $R_{ij}$  = the proportional responsibility that  
859 country j has for a status change in species i, where for each species i:

860

$$861 \quad \sum_j R_{ij} = 1.0$$

862

863

864 For brevity, we use the phrase “proportional responsibility” (or simply  
865 “responsibility”) to refer to the relative influence that factors in each country had on  
866 the changing conservation status of each species. Proportional responsibilities cannot  
867 be known exactly, and so the algorithm will generate estimates of responsibility with  
868 some error. For predictive modelling, an equally important condition of algorithm  
869 design is that such errors should not bias regression outcomes.

870 The most commonly used responsibility algorithm simply counts the number of  
871 declining species in each country (usually, the number of species classified as having  
872 some level of threat in global Red List assessments)<sup>14,15,45,46,77</sup>. Implicitly, such an  
873 algorithm assumes that if two countries share a species, they have equal responsibility  
874 for that species' decline. This is reasonable if both countries have roughly equal  
875 shares of the species range. However, species are frequently distributed so that one  
876 country holds the bulk of the range (e.g. >80% of the range) and neighbouring  
877 countries hold very small fractions of the remaining range edge (e.g. <5% each)

878 (Extended Data Figure 5). In such cases, it would be highly inaccurate (and politically  
879 unfair) to allocate equal shares of responsibility for a species decline across all these  
880 countries. A fairer, more accurate system may be to divide up responsibility according  
881 to the fraction of each species' range found in each country<sup>8,26</sup>. Formally, if  $p_{ij}$  is the  
882 proportion of the range of species  $i$  in country  $j$ , then the value of  $p_{ij}$  is an estimate of  
883 the true responsibility  $R_{ij}$ , with some error implied in that estimate (formally, the error  
884 is defined as the difference between the  $p_{ij}$ -based estimate and  $R_{ij}$ ).

885 For any observed  $p_{ij}$ , there is therefore a theoretical probability density function  
886 (PDF) of all possible  $R_{ij}$  that it could represent. For example, if a species is split 60:40  
887 between two countries, then for the  $p_{ij} = 0.60$  country, the underlying assumption is  
888 that there is an approximately Gaussian PDF for  $R_{ij}$  with a central mode at 0.6, such  
889 that the most probable value of  $R_{ij}$  is 0.60 or close to it, whereas extreme values such  
890 as 0.0 or 1.0 have a very low theoretical probability.

891 First imagine that for any country  $j$ , all  $p_{ij} = 0.60$ , and so all  $R_{ij}$  follow a Gaussian  
892 distribution around 0.6. The range-based algorithm will generate a series of positive  
893 and negative errors  $eR_{ij}$  (=overestimates and underestimates of  $R_{ij}$ ). The same is true  
894 of the country with  $p_{ij} = 0.40$ . However, the true quantity of interest we wish to  
895 estimate is  $BDS_j$  (i.e. the sum of  $R_{ij}$  rather than each individual  $R_{ij}$ ). There is therefore  
896 an associated set of errors

$$898 \quad eBDS_j = \left( \sum_i eR_{ij} \right)$$

899

900  
901 For a predictive regression model, the critical question is whether these errors  
902  $eBDS_j$  are likely to strongly affect modelling of  $BDS$ , for example by creating  
903 artefactual patterns or biased, non-random error distributions. If all range splits that  
904 make up  $BDS_j$  are relatively symmetric (i.e. similar to 60:40), then it is a reasonable  
905 expectation that the errors, being drawn from an approximately Gaussian distribution,  
906 will overestimate and underestimate with relatively equal frequency, and so the sum  
907 of errors will not depart strongly from zero. Thus, the errors are expected to be  
908 relatively random in their distribution, permitting robust modelling. It is also  
909 particularly unlikely that the errors would create artefactual impacts, since this would  
910 require a consistent, non-random association between large negative errors and  
911 higher-spending countries (sufficiently large, indeed, to strongly depress  $BDS_j$ ), plus  
912 equally large and consistently positive errors for lower-spending countries.

913 However, when  $p_{ij}$  is closer to its limits of 0.0 and 1.0, biased errors become  
914 highly likely. Human-induced population losses (leading to species declines and Red  
915 List status changes) are generally focused spatially in the particular part or parts of the  
916 species range where human pressures have most strongly increased and in general, it  
917 is very rare for such hotspots of decline to lie around the range periphery<sup>36</sup>. Therefore,  
918 a country that holds 3% of the species range will often have zero responsibility rather  
919 than 3% responsibility, and the neighbour with 97% of the range will often be entirely  
920 responsible for a status change. Even in a random process (with limited trials and  
921 therefore stochastic outcomes), spatial clusters of increased mortality dropped at  
922 random onto the range will frequently fall entirely within the 97% country. Formally,  
923 therefore, when  $p_{ij} = 0.03$ , the associated probability density for  $R_{ij}$  will be high at 0  
924 and decline rapidly towards a very low density at  $R_{ij} = 0.03$ , giving a PDF with a  
925 strong right skew and a likely 99<sup>th</sup> percentile at around  $p_{ij}$  itself.

926 In the example where  $p_{ij} = 0.03$ , therefore, nearly all errors will be overestimates,  
 927 and the most common likely scenario is an overestimate of exactly 0.03. Generalising,  
 928 whenever  $p_{ij}$  is small and the PDF is right skewed, a raw or “unadjusted” range-based  
 929 algorithm will overestimate responsibility in almost all cases, generating highly  
 930 biased errors  $eR_{ij}$  that will commonly have magnitude  $+p_{ij}$ . By the same process, using  
 931  $p_{ij}$  to estimate  $R_{ij}$  at high  $p_{ij}$ , such as 0.9, will tend to underestimate true responsibility  
 932 in the great majority of cases.

933 The critical question is how severely this consistent bias will affect the  
 934 regression analysis. We examined the data and found that empirically, a large number  
 935 of countries had a BDS composed entirely of a trivially small (e.g. <5%) range edge  
 936 fractions (Extended Data Figure 5). Their  $BDS_j$  estimates were therefore likely to be  
 937 made up of multiple small  $p_{ij}$  that were consistently overestimating responsibility  $R_{ij}$ .  
 938 In analysing  $BDS_j$ , the error metric of interest is  $eBDS_j =$  the sum of  $eR_{ij}$ . Since the set  
 939 of errors  $eR_{ij}$  was likely to be highly biased and the most common likely scenario was  
 940 that  $eR_{ij} = +p_{ij}$ , then  $eBDS_j$  (as the sum of  $eR_{ij}$ ) would also be highly biased, with a  
 941 substantial probability that  $eBDS_j$  would equal sum ( $p_{ij}$ ). Since all the individual  $p_{ij}$   
 942 values comprising these BDS scores were both trivially small and likely overestimates  
 943 of zero, the associated BDS scores were also likely to be trivially small (and biased)  
 944 overestimates of zero. We refer to these cases as range-edge BDS or “reBDS scores”.

945 We further explored the empirical impact of this suspected bias on the  
 946 information signal by making exploratory plots of BDS against its possible predictors.  
 947 These plots showed that reBDS scores indeed generated a dense cloud of very small  
 948 values, close to the x axis, that was visually distinct from patterns across larger (and  
 949 likely more accurate) BDS. In Signal Detection Theory terms<sup>37</sup>, therefore, reBDS  
 950 cases were highly likely to represent strong signal noise that also lay non-randomly to  
 951 one side of the main information pattern, in a cloud of such density that the signal-to-  
 952 noise ratio was extremely low, the ability of regression models to detect predictive  
 953 relationships was compromised, and any calculated model parameters were likely to  
 954 be strongly biased by the non-random error. Similarly, in the binomial analysis, the  
 955 same reBDS issue caused many species-poor countries to have  $BDS_b = 1$  purely  
 956 because those countries contained trivial range edges of status-changing species found  
 957 almost entirely elsewhere.

958 To reduce these issues of signal noise and bias at small  $p_{ij}$ , we explored setting  
 959  $R_{ij}$  to zero for small  $p_{ij}$ . Formally, we explored setting a threshold value  $T$ , such that  
 960 responsibility was set to zero for any country with a range fraction  $< T$ , such that

961

$$962 \quad R_{ij} = \quad p^*_{ij} = \quad \begin{cases} p_{ij} & \text{if } p_{ij} \geq T \\ 0 & \text{if } p_{ij} < T \end{cases}$$

963

964 (but see below for  $p_{ij} \geq (1-T)$ ).

965

966 To decide on appropriate values for the threshold  $T$ , we used Signal Detection  
 967 Theory in combination with the mathematics of the Red List criteria. The most  
 968 important aspect of this approach that when  $p_{ij}$  is small (e.g. 0.03), true  $R_{ij}$  may often  
 969 but not always be zero, but it is impossible to know which range-edge countries  
 970 genuinely had a very small responsibility, and which had a true-zero responsibility.  
 971 Therefore, reBDS values will often but not always be non-zero overestimates of a true  
 972 zero. In Signal Detection Theory, cases where a true zero is wrongly given a non-zero  
 973 value represent “false positives”. However, any threshold could also cause the  
 974 algorithm to wrongly exclude (set to zero) some cases where the reBDS score  
 975 represented a genuine (if small) fractional responsibility, and such incorrect

976 exclusions are classed as “false negatives”. The higher the threshold  $T$ , the more false  
977 positives will be correctly excluded but the more false negatives will be wrongly  
978 excluded. Theoretical optimisation will therefore seek values of  $T$  large enough to  
979 avoid too many false positives (i.e. guarding against picking up too much noise) yet  
980 small enough to avoid too many false negatives (i.e. guarding against throwing away  
981 too much information). A threshold that produces too many false positives is classed  
982 as overly “sensitive” and one that produces too many false negatives is classed as  
983 overly “specific”.

984 For BDS, the optimal signal detection threshold cannot be precisely estimated  
985 because the proportions of false positives and false negatives at any value of  $T$  are not  
986 empirically known, and so the ratio of sensitivity to specificity cannot be calculated.  
987 Appropriate thresholds therefore need to be estimated by theoretically estimating the  
988 optimal sensitivity/specificity trade off. Furthermore, in this analysis, sensitivity and  
989 specificity were likely to have different impacts on analytical bias and outcomes  
990 (making approaches that give equal weight to sensitivity and specificity, or that  
991 require accurate knowledge of the ratio between them e.g. area under the curve<sup>37</sup>, less  
992 appropriate). The main deleterious effect of excessive sensitivity was to generate large  
993 amounts of biased noise, as already shown. The main impacts of excessive specificity,  
994 on the other hand, were likely to be (a) to slightly underestimate BDS (because a few  
995 small responsibility fractions had been wrongly discarded); (b) to reduce sample size  
996 for the continuous model part (because of removing reBDS countries); and (c) to  
997 change the ratio of ones to zeroes in the binomial analysis (because reBDS countries  
998 have  $BDS > 0$  before adjustment and  $BDS = 0$  after adjustment). Since high levels of  
999 noise and bias associated with lack of specificity are likely to have a much stronger  
1000 impact than the small underestimates and sample size/binomial ratio effects associated  
1001 with lack of sensitivity, avoiding false positives should take priority.

1002 To allocate this priority (i.e. to avoid repeatedly replacing true zeroes with  
1003 trivially small values), the algorithm needs to set  $T$  such that for all probability  
1004 frequency distributions associated with all  $p^*_{ij}$ , there is a low probability density at  $R_{ij}$   
1005  $= 0.0$ . Formally, we set a target that for all  $p^*_{ij}$ ,  $\text{prob}(R_{ij} = 0.0)$  should be  $< 0.5$  and  
1006 ideally  $\ll 0.5$ . However, a second consideration is that in range-edge countries, the  
1007 likely probability density at zero is affected by the size of decline implied by a status  
1008 change. To illustrate this, we take the example of a country that holds 10% of a  
1009 species’ range and the most frequent criterion justifying a status change, population  
1010 loss (Red List category A(2-4)<sup>3</sup>). When population loss occurs, the Red List  
1011 assessment for any particular period is based on a rate of change over time, and so a  
1012 change in Red List status expresses a second-derivative change in the rate of change  
1013 i.e. additional net mortality/disappearance over and above what had occurred in the  
1014 previous assessment period. Clearly, if a status change formally represented a 99%  
1015 increase in mortality/disappearance for the entire species, there would be a strong  
1016 probability that at least some of those additional deaths or disappearances had  
1017 occurred in the 10%-holding country. However, genuine status changes generally  
1018 imply an increase in loss of a few tens of percentage points. For example, a common  
1019 status change is LC to VU, where LC can imply anywhere between zero decline and  
1020 29.9% loss over a period of ten years or three generations, and VU is defined as  
1021 anywhere between 30% and 49.9% loss (depending on the use of the near-threatened  
1022 category by assessors)<sup>81</sup>. If we take the midpoints of these ranges (15% and 40%  
1023 respectively), then an LC-to-VU change would indicate an average 25 percentage  
1024 point increase in loss (the difference between 40% and 15%), while other changes not  
1025 at the exact midpoints would indicate a difference in decline rates above or below 25.

1026 Since the additional deaths underlying a status change are generally non-  
1027 randomly clustered in geographic space as wave fronts expanding from points of  
1028 increased human pressure<sup>36</sup>, this 25-point change can be imagined as a small number  
1029 of clusters of additional net loss placed onto a gridded range, where the 10%-holding  
1030 country occupies the leftmost 10% of the grid and another country or countries, the  
1031 rightmost 90%. Often, such spatially-clustered mortality increases might be expected  
1032 to fall entirely within the rightmost 90%, implying that a 10%-holding country will  
1033 frequently have no responsibility. To explore this intuition this more quantitatively,  
1034 we simulated a 25-point population loss as a varying (2-5) number of rectangular  
1035 blocks that covered a total of 25% of a 10x10 gridded range. The first column of the  
1036 grid was then treated as the 10%-holding country and the remaining 9 columns to  
1037 another country or countries: (it is moot whether it is one or several countries in the 9  
1038 columns because the simulation focuses only on the likelihood that the 10% country  
1039 will not have any part of any decline cluster overlapping its territory). The blocks  
1040 were then placed independently of each other, for a limited number of trials (n=100)  
1041 to introduce stochasticity, onto the gridded range and for each placement, we tested  
1042 whether any part of the leftmost column had been overlapped. Overall, we found that  
1043 the probability of any overlap between a block and the leftmost 10% of the grid was  
1044 generally <0.5, varying with the number of blocks. For example: if the decline occurs  
1045 as two independently-placed blocks, the simulated probability of overlap was 0.19,  
1046 giving a 0.81 probability that the range-edge country has  $R_{ij} = 0$  (i.e. an 81% chance  
1047 of a false positive). When the 25-point decline was modelled as five independently  
1048 dropped blocks, the overlap probability rose to 0.41, indicating a 59% chance of a  
1049 false positive – still appreciably greater than our target false-positive rate of <<0.50.  
1050 These values are also conservative because clusters of loss are often not spatially  
1051 independent of each other but rather, may be grouped due to larger-scale spatial  
1052 contagion in threats and associated losses<sup>36</sup>. Such grouping further reduces the  
1053 random probability of an overlap with the range edge and thus, would increase the  
1054 false positive rate further. Similar outcomes occur for other percentage point increases  
1055 in mortality, as implied by other IUCN status changes.

1056 Indeed, even if the 25-point population loss is unrealistically (and highly  
1057 conservatively) modelled as spatially homogeneous, then define  $q$  = the change in rate  
1058 of species decline required to trigger a change in Red List status (such that in the  
1059 example,  $q = 0.25$ ). Under an assumption of homogeneity, the theoretical maximum  
1060 responsibility that a 10%-holding country can have for 25% change is ~40%, or  
1061 10/25. More formally, we define the 99<sup>th</sup> percentile of theoretically probable  $R_{ij}$  for  
1062 the 10%-holding country as  $p_{ij}/q = 0.1/0.25 = 0.4$ . A distribution with a 99<sup>th</sup> percentile  
1063 at 0.4 is likely to have a relatively strong skew and consequently, a relatively high  
1064 probability density at  $R_{ij} = 0$ , since skewness in the theoretical probability distribution  
1065 for  $R_{ij}$  increases at an accelerating rate as the entire distribution moves to the left.

1066 There is therefore a strong likelihood that even for non-trivial  $p_{ij}$ , such as 10% or  
1067 more, the probability that ( $R_{ij} = 0$ ) will be greater than the algorithm's target of <<0.5.  
1068 Therefore, the theoretical expectation is that to avoid false positives to a sufficient  
1069 degree, the threshold  $T$  may need to be set at greater than 0.1 and potentially as high  
1070 as 0.2 or more. To further explore this expectation empirically, we further examined  
1071 exploratory biplots of BDS against its predictors when  $T$  is varied between 0.05 and  
1072 0.25. We found that as  $T$  was reduced, and as expected from our theoretical treatment,  
1073 increasingly large numbers of likely false positives became included in the BDS  
1074 dataset, with noise increasing rapidly at  $T < 0.1$  (i.e. an increasingly dense cloud of  
1075 points with trivially small BDS values developed). On the other hand, increasing  $T$



1076 from 0.14 to 0.25 caused little variation in  $R_{ij}$  values themselves, but progressively  
 1077 reduced sample size (and so power) in the continuous analysis, with the drop off in  
 1078 sample size being small between  $T = 0.1$  and  $T=0.17$ , then larger between  $T = 0.17$   
 1079 and  $T = 0.25$  (see Supplementary Results).

1080 Simulation and probability theory can therefore suggest the approximate range  
 1081 for appropriate values of  $T$  but the exact optimal value must remain uncertain. To  
 1082 account for this uncertainty and its possible impact on model outcomes, we performed  
 1083 our final analysis three times for three different values of  $T$ : 0.10, 0.17 and an extreme  
 1084 value of 0.25. The main text of the paper shows results for  $T = 0.17$ , being the  
 1085 parameter value where false positives could be reduced as far as possible, and yet  
 1086 without the trade-off of sample size reduction becoming severe; results for  $T = 0.10$   
 1087 and  $T = 0.25$  are described in Extended Data Table 4 and Supplementary Results.

1088 In formal summary, for each species  $j$ , each country  $i$  holds  $R$  proportional  
 1089 responsibility for the total decline  $d$  of  $j$ . Decline  $d$  can be positive and indicate a  
 1090 worsening extinction risk ( $d>0$ ), it can be negative and indicate a reduction in  
 1091 extinction risk (“negative decline” i.e. an improvement,  $d<0$ ) or it can be constant  
 1092 ( $d=0$ ). Each country’s baseline Biodiversity Decline Score ( $BDS_i$ ) is therefore the net  
 1093 sum of all its decline fractions and improvements (negative decline fractions):

$$1094 \quad BDS_i = \sum_j d_j R_{ij} \quad (1)$$

1095

1096 where

$$1097 \quad R_{ij} = p^*_{ij} / \sum_i p^*_{ij} \quad (2)$$

1098 where  $p^*$  indicates the range proportion of each species  $j$  in country  $i$  after range  
 1099 fractions below the minimum percentage  $T$  have been set to zero, or formally:

$$1100 \quad p^* = \begin{cases} p_{obs} & \text{where } p_{obs} \geq T \\ 0 & \text{where } p_{obs} < T \end{cases}$$

1101  
 1102 where  $p^*$  indicates the range proportion of each species  $j$  in country  $i$  after range  
 1103 fractions below the minimum percentage  $T$  have been set to zero, or formally:  
 1104  
 1105  $p^* = \begin{cases} p_{obs} & \text{where } p_{obs} \geq T \\ 0 & \text{where } p_{obs} < T \end{cases}$   
 1106  
 1107  
 1108 If a species is split 95:5 between two countries and the responsibility  $R$  has been  
 1109 set to zero for the 5%-holding country, then for consistency,  $R$  for the 95%-holding  
 1110 country should be increased from 0.95 to 1.0, and equation (2) indeed performs this  
 1111 function. However, a widespread species can be spread in small fractions across  
 1112 multiple countries without any one country having a major proportion of the range. In  
 1113 such cases, if only one country has a range fraction exceeding the threshold (e.g.  
 1114 17.1%) then under equation (2), that country would receive a clearly exaggerated  
 1115 100% of responsibility for the change in risk status (whereby  $p_{obs} = 0.171$  but  $p^* =$   
 1116 1.0). Such cases as fairly rare (widespread species rarely move out of the Least  
 1117 Concern category) but to avoid any such error, we reset the denominator of equation  
 1118 (2) to unity whenever a widespread species was scattered in small fractions across  
 1119 multiple countries.

1120 To calculate the  $p_{ij}$  fractions themselves, we extracted the percentage of the  
 1121 geographic range of all global bird and mammal species contained within the national  
 1122 borders of each country (the range overlap)<sup>26</sup>. Range overlap for mammals was  
 1123 extracted using ArcGIS utilities on the range maps provided by the IUCN Global

1124 Mammal Assessment<sup>82</sup> (see<sup>8</sup> for details). This procedure gave very exact areas of  
1125 overlap for the taxon Mammalia, but the calculation required us to run twenty  
1126 processors in parallel for nearly a month. For the much larger taxon Aves, therefore,  
1127 we used a slightly different procedure. Bird ranges were obtained as polygons in ESRI  
1128 shapefiles provided by Birdlife International<sup>83</sup>. Species range areas that were  
1129 designated as non-native or dubious presence were excluded *a priori*. For each  
1130 species X, we then combined wintering and breeding ranges (because threats to bird  
1131 species can occur in both their breeding and wintering ranges), and gridded all range  
1132 polygons onto a 0.1 degree raster grid, using a cylindrical equal-area projection to  
1133 match the projection of the original vector data. We designated all grid cells that had a  
1134 center point lying inside a range polygon for X as 'presence cells' for X, overlaid each  
1135 presence cell onto a vector dataset of the world's countries<sup>84</sup> using the `over` and  
1136 `wrld_simpl` functions in R packages 'sp'<sup>85</sup> and 'maptools'<sup>86</sup>, allocated the cell to the  
1137 country found at the cell centre point, and then calculated the fraction of all presence  
1138 cells for X found in each country. Prior to this calculation all countries with coastlines  
1139 were enlarged by a 0.05 degree buffer into the sea to account for responsibility of sea  
1140 bird ranges in coastal waters; coastal marine mammals were treated in a similar way,  
1141 see<sup>8</sup>.

1142 As an additional accuracy check, we examined individual Red List reports for  
1143 every declining species to see where the range-based approximation of responsibilities  
1144 was clearly inappropriate, and revised those cases accordingly. Our revisions are  
1145 listed in Supplementary Table 4 and include cases where (i) a decline had majorly  
1146 affected how the geographic range was distributed across countries, including cases  
1147 where a species had once been found in other countries but was now missing from  
1148 them; (ii) the species population distribution across countries was poorly correlated  
1149 with the range distribution; and (iii) specified actions e.g. along migratory routes had  
1150 an impact clearly disproportionate to the percentage of the global range found in the  
1151 country carrying out those actions.

1152 At a theoretical extreme, a 100% range fraction for a declining species could  
1153 indicate that one country contains the last extant individuals of a species that used to  
1154 be widespread in neighboring countries. The 100%-holding country would then  
1155 represent a final "oasis" at the species' former range edge, and it would be wholly  
1156 unjust to assign 100% responsibility for the decline to it. However, our assumption is  
1157 that in the mere eight-to-twelve years between our IUCN assessments, there will  
1158 rarely be a case where a species has been extirpated from its main homeland countries  
1159 without some record of this event existing. We applied the BDS adjustments based on  
1160 Red List reports after the adjustments for range edges (reBDS), and so our method  
1161 corrected for any such anomaly. For example, *Addax nomasculatus* (the rare  
1162 screwhorn antelope) has recently disappeared from Chad and Mali and so we  
1163 incremented the BDS of those two countries to reflect this (Supplementary Table 4).

1164  
1165

#### 1166 Data availability

1167 The authors declare that the data supporting the findings of this study are  
1168 available within the Supplementary Information; original socioeconomic data (except  
1169 governance values) can also be sourced from the World Bank  
1170 <http://databank.worldbank.org>; original governance values can also be sourced from  
1171 the Worldwide Governance Indicators dataset [www.govindicators.org](http://www.govindicators.org) (governance  
1172 data).  
1173

1174            Code availability  
1175            R scripts used in analysis are available upon request from the corresponding  
1176 author.  
1177  
1178

1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237

## EXTENDED DATA LEGENDS

**Extended Data Figure 1 | The average per-species BDS for each country (i.e. BDS/total fractional species richness, expressed as a percentage).** Dark red = > 5%, dark orange/red = 2.5 – 5%, mid orange = 1 – 2.5%, pale yellow = 0 – 1%, grey = 0%, blue = improving (negative percentage), light grey hatching = cannot be calculated (zeroes in the denominator). Note that in more species-poor countries e.g. much of Europe and the Arab geographic crescent, zeroes are expected at random (supplementary methods). See Supplementary Table 1 for precise values per country. Country outlines supplied by esri\_dm

<https://www.arcgis.com/home/item.html?id=d86e32ea12a64727b9e94d6f820123a2#overview>

**Extended Data Figure 2 | The effect of agricultural expansion on decline (binomial part, n=109 independent countries) depends on both governance improvement and the existing percentage of land converted.**

The effect (coefficient) of agricultural expansion on the probability of a decline occurring is shown on the y axis and varies with the rate of governance improvement on the x axis. Coefficients >0 (above the dashed line) indicate that agricultural growth increases the probability of a decline occurring, v.v. for <0. However, the coefficient further depends on a second moderator, the % of land converted to agriculture: red =50<sup>ilc</sup> of % land conversion, grey =25<sup>ilc</sup>; lines show mean and coloured bands show conditional 95% confidence intervals. Note how effects are most strongly deleterious on less heavily converted landbases. Rug plot at bottom shows empirical distribution of x-axis values (but note that countries with more % agric. land generally have slow governance improvement). All variables are z-standardised.

**Extended Data Figure 3 | The impact of conservation spending on decline depends on threatened species richness and on GDP.** (a) Spending effect size and threatened species richness (continuous part, n=50 independent countries); (b) spending effect size and GDP (binomial part, n=109 independent countries). The effect size (coefficient) for spending is shown on the y axis and varies with the value of species richness on the x axis. The more negative the coefficient is on the y axis, the more strongly spending reduces declines (continuous) or the probability of a decline occurring (binomial). Conditional confidence bands are shown; rug plots at bottom show empirical distribution of x-axis values. All variables are z-standardised.

**Extended Data Figure 4 | Observed declines versus model-predicted declines.** (a) BDS versus predicted BDS in the continuous part (n=50 independent countries). Both axes are ln-transformed for clarity; (b) As (a), but zooming in to the lower-BDS countries only (note axes values in (a) and (b)); (c) Observed decline events (BDSb) versus the predicted probabilities of a decline event, from the binomial part (n=109 independent countries). Observed decline events on the x axis (0 = no decline occurred, 1 = decline occurred) have been jittered for visibility; (d) Change in model prediction when top 3 BDS values are excluded: black line = full dataset prediction, dashed red line = prediction with exclusions.

**Extended Data Figure 5 | Distributions of BDS and species range fractions across countries.** (a) Index plot of BDS scores. For clarity, BDS has been ln(x+10) transformed, and so the straight line at 2.3 shows the long tail of zeroes. (b) Distribution of all range fractions in all countries, showing the very large number of small, range-edge fractions (<10% of a species is found in a country). (c) Distribution of the maximum range fraction for all species, showing how a large number of species have >90% of their range in one country. (d) Distribution of the minimum range fraction for all species, showing how very many species have a small range edge (<10% of their range) in a second country.

**Extended Data Figure 6 | Differences in absolute Pearson's correlations between conservation spending and each of its covariates before and after carrying out covariate balancing propensity score weighting (CBPS).** (a) continuous analysis; (b) binomial analysis. Upper bars show absolute Pearson correlations prior to CBPS, lower bars after CBPS. Box shows interquartile range with the median (bold central line). Whiskers show most extreme data point no more than 1.5 times the interquartile range. N=50 independent countries.

1238

1239

1240

1241

1242

1243

**Extended Data Table 1 | List of regression terms tested. Also shown are the best-fitting four models from continuous analysis with their AICc values, Akaike weights and variables (see Supplementary Table 2 for full continuous-model results).** Spending = conservation spending PPP; Agric. = agricultural; governance = government effectiveness indicator. In main body of table, 1 = term included, 0 = term not included.

1244

1245

1246

1247

1248

**Extended Data Table 2 | Cross correlations between variables.** \$\$ = conservation spending PPP; Agric. = agricultural; Pop = population; Gov = governance; Decl = declines; Spp. Rich = threatened species richness; For. Loss = % forest loss; Area = country area.

1249

1250

1251

1252

**Extended Data Table 3 | Variance inflation factors (VIFs) for the continuous and binomial model parts.** Spending = conservation spending PPP; Agric. = agricultural; Pop = population; Gov = governance; Spp. Rich = threatened species richness; Area = country area.

1253

1254

1255

1256

1257

1258

1259

1260

1261

1262

1263

1264

**Extended Data Table 4 | Standardized coefficients for best-fitting models under alternative assumptions.** Best-fit models that used alternative values of the threshold T are shown. We very strongly caution that for interacting variables (marked \*), the coefficients shown cannot be interpreted by simply reading the table (refer to the Supplementary Results for their complex interpretation). “Agric. land” = mean percentage of agricultural land; t-1 = 1994-2000, t-2 = 1988-1994; GDP = Gross domestic product per capita PPP. Population = rural population density; governance improvement = change in the government effectiveness score. For T=0.10, sample size increased to n=53 independent countries in the continuous part (index parameter = 1.99), and the ratio of ones to zeroes was 44:65 in the binomial part. Equivalent values for T=0.25 are n=43 independent countries (i.e. a large sample size decrease) and a ratio of 37:74.





