**Supplementary Material 4.** Alternative statistical analysis

We chose to use partial least squares structural equation modeling (PLS-SEM) instead of traditional panel analysis approaches to analyze these data for two of reasons. First, PLS-SEM can account for relationships among drivers (e.g., between GNI and healthcare spending, between biodiversity and forestation), while such collinearity violates model assumptions and can result in invalid coefficient estimates in panel analysis. Second, whereas panel analysis seeks explanatory factors that are consistent in space and time, we were interested in some processes that play out differently across space than they do through time. For instance, biodiversity is higher in warm tropical countries for a variety of evolutionary reasons, including greater availability of energy and greater time available for speciation. But climate warming is expected to decrease biodiversity because it will shift existing environmental conditions outside of the thermal range for many species. Although we chose to focus on the results of the PLS-SEM analysis in the main text, our data were suitable for a panel (general linear mixed model or GLMM) analysis parallel to the SEM presented in the main text, once we accounted for collinearity by monitoring variance inflation factors (VIF) and dropping any highly collinear (VIF >10) factors. This parallel approach allows us to test the robustness of the PLS-SEM approach with a separate, independent analysis.

We created a single GLMM model for the burden (annual global DALYs) of each disease. Models included fixed effects of forestation, biodiversity (mammals + birds), population density, GNI, urbanization, precipitation, temperature, whether the country is in New or Old World, and year (1990 or 2010), along with a random effect of country (to control for the fact that countries are represented at two time points; see Supplementary Material 2 for full description of all variables). Two sets of these models were created – one for forestation as measured by satellite, and another for forestation as reported to the FAO (see Supplementary Material 2).

We then used meta-analysis on the standardized regression coefficients from GLMM models to detect general patterns among diseases. We derived standard errors for each standardized regression coefficient using the standard deviation associated with the SEM model divided by the square root of the number of countries for which that model was run (Table 1).We calculated a cumulative effect size for each driver across all diseases, using a random-effects model weighted by the inverse of the variance for each effect size. All meta-analyses were performed in the *metafor* package in R.

Results from these GLMMs were strongly consistent between the models that used satellite-derived forest data and the models that used forest data as reported to the FAO (coefficient = 1.0306, intercept = 0.0000; z = 4.6276, df = 5, p < 0.0001; Figures S1 and S2), and were also consistent with results from the PLS-SEM presented in the main text. Just as the SEM approach taken in the main text showed, the GLMM approach revealed a strong negative association between urbanization and disease burden, and a positive association between temperature and disease burden (Figure S1). As in the PLS-SEM, the GLMMs demonstrated that wealth (here, GNI) had a non-significant negative association with disease burden and that biodiversity had a non-significant positive association with disease burden. In both models, population, forestation, and precipitation were not strongly associated with disease burden.

The concordance of the PLS-SEM and GLMM approaches lends credibility to our conclusions. These results are consistent across models that differ significantly in their assumptions, specifications, and limitations, giving us confidence in their robustness.



**Figure S1**. Results of meta-analyses for models using satellite-derived forest cover (x-axis) and forest cover reported to the FAO (y-axis), which summarize the results of GLMMs across 24 diseases. Points represent mean effect sizes for the effects of % of population living in urban environments, forest cover, and temperature. Error bars represent 95% confidence intervals. Other factors clustered around 0 in this plot are shown in Figure S2 (bounded by red inset box shown here).



**Figure S2**. Results of meta-analyses for models using satellite-derived forest cover (x-axis) and forest cover reported to the FAO (y-axis), which summarize the results of GLMMs across 24 diseases. Points represent mean effect sizes for the effects of % of population living in urban environments, forest cover, and temperature. Error bars represent 95% confidence intervals. This plot is bounded by the red inset box shown in Figure S1.