The non-independence of nations and why it matters

Scott Claessens[1] & Quentin D. Atkinson[1,2]

[1] School of Psychology, University of Auckland, Auckland, New Zealand

[2] Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Author Note

This working paper has not yet been peer-reviewed.

The authors made the following contributions. Scott Claessens: Data curation, Formal analysis, Investigation, Software, Visualization, Writing - original draft, Writing - review & editing; Quentin D. Atkinson: Conceptualization, Funding acquisition, Supervision, Writing - review & editing.

Correspondence concerning this article should be addressed to Quentin D. Atkinson, Floor 2, Building 302, 23 Symonds Street, Auckland, 1010, New Zealand. E-mail: q.atkinson@auckland.ac.nz

Abstract

Cross-national analyses test hypotheses about the drivers of global variation in national outcomes. However, since nations are connected in various ways, such as via spatial proximity and shared cultural ancestry, cross-national analyses often violate assumptions of non-independence, inflating false positive rates. Here, we show that, despite being recognised as an important statistical pitfall for over 200 years, cross-national research in economics and psychology still does not sufficiently account for non-independence. In a review of the 100 highest-cited cross-national studies of economic development and values, we find that controls for non-independence are rare. When studies do control for non-independence, our simulations suggest that most commonly used methods are insufficient for reducing false positives in non-independent data. In reanalyses of twelve previous cross-national correlations, half of the effect sizes show no association after controlling for non-independence using global proximity matrices. We urge social scientists to sufficiently control for non-independence in cross-national research.

Keywords: spatial non-independence; cultural non-independence; cross-national analyses; simulations; replications; reanalyses

Word count: xxxx words

The non-independence of nations and why it matters

Nations are perhaps the single most important population unit structuring modern human life. The nation in which someone is born has a large effect on what they can expect out of life, including their income level[1], life expectancy[2], mental health[3], subjective well-being[4], and educational attainment[5]. Nations are also among the most important determinants of human cultural variation, with greater cultural similarity within than outside national borders[6].

Given the importance of nations for structuring human behaviour, there is justifiably huge interest in statistical analyses that attempt to predict variation in national outcomes of all kinds. At the time of writing, a search in Web of Science for the term "cross-national" in titles or abstracts returned over 13,000 unique hits. The standard practice for cross-national analyses is to conduct bivariate correlations or multiple regressions with individual data points representing different nations. Such analyses widen the scope of social science beyond Western populations[7,8] and have been used to study, among other topics, the causes of variation in the economic wealth of nations[9–12], global patterings of cultural norms and values[13–16], and the universality and diversity of human behaviour and psychology around the world[17–20].

However, cross-national analyses are complicated by the fact that nations are not statistically independent data points. Unlike independent random samples from a population, nations are related to one another in a number of ways. First, nations that are closer to one another tend to be more similar than distant nations. This phenomenon is known as spatial non-independence[21], and it occurs because nations in close spatial proximity share characteristics due to local cultural diffusion of ideas[22] and regional variation in climate and environment[21]. For example, the neighbouring African nations Zambia and Tanzania have similar levels of terrain ruggedness, which has been used to partially explain their similar levels of economic development[23]. This pattern conforms to

Tobler's first law of geography: "everything is related to everything else, but near things are more related than distant things"[24] (p. 236).

Second, nations with shared cultural ancestry tend to be more similar than culturally unrelated nations. This is known as cultural phylogenetic non-independence[25–27], and occurs because related nations share cultural traits inherited via descent from a common ancestor. Shared cultural ancestry can result in a form of pseudoreplication, whereby multiple instances of the same trait across nations are merely duplicates of the ancestral original. For example, the related island nations Tonga and Tuvalu share similar languages and customs due to cultural inheritance from a common Polynesian population dating back more than 1,000 years. More recently, shared ancestry explains a myriad of cultural similarities between colonial settlements and their colonisers (e.g. Argentina and Spain). Importantly, these deep cultural connections between nations often span large geographic distances around the world. Tonga and Tuvalu share cultural traits despite being separated by over 1,500 kilometres of ocean, and Argentina and Spain remain culturally similar today despite their location on two separate continents. Shared cultural ancestry must therefore be considered independently of spatial proximity in the study of nations.

Spatial and cultural phylogenetic non-independence between nations make cross-national inference challenging. A fundamental assumption of regression analysis is that model residuals should be independently and identically distributed[28]. But without accounting for spatial or cultural non-independence between nations, model residuals can show structure that remains unaccounted for, violating this assumption. Treating nations as independent can thus inflate false positive rates[29], producing spurious "direct" relationships between variables that in fact only indirectly covary due to spatial or cultural connections[30] (see Supplementary Figure S1 for an example causal model).

Non-independence is widely acknowledged in fields that routinely deal with spatially or culturally structured data. In ecology and sociology, studies with regional-level data use a variety of autoregressive models designed to account for spatial patternings[31,32]. In anthropology, researchers have recognised cultural non-independence as an important statistical pitfall for over 200 years, with issues of cultural pseudoreplication being identified in early comparative studies of marriage practices across societies[25]. In the twentieth century, anthropologists began to emphasise that human societies do not develop independently, but rather exist in a globally interconnected "world system" linked by shared history and cultural ancestry[33]. Researchers compiled the Standard Cross-Cultural Sample of 186 cultures to minimise the confounding effects of this non-independence in comparisons of human societies[34], though spatial and cultural dependencies are difficult to remove entirely[35,36]. Today, anthropologists borrow phylogenetic comparative methods from evolutionary biology, such as phylogenetic least squares regression[37], when comparing societies, treating culturally related societies in the same way as biologists treat genetically related species (e.g.[38,39]).

At the national level, recent reanalyses have revealed that several cross-national relationships reported in economics and psychology do not hold when controlling for non-independence between nations. One working paper replicated 25 analyses of "persistence" in economics, in which modern national outcomes are regressed against historical characteristics of those nations, and found that over half of the relationships were attenuated when controlling for spatial non-independence[40]. Another replication study found that many of the widely publicised relationships between national-level pathogen prevalence and political institutions and attitudes fail to hold when controlling for various kinds of non-independence[41]. These reanalyses, and others[42–44], raise the question: how widespread a concern is non-independence in studies of national-level outcomes?

To address this question, we consider national-level variables of general interest across the social sciences: economic development and cultural values. These variables are frequently included as both outcomes and predictors in cross-national studies in economics and psychology[9–16]. First, we show that economic development and cultural values are spatially and culturally non-independent across nations, emphasising the need to control for non-independence. Second, we systematically review the 100 highest-cited cross-national studies of economic development and cultural values and estimate the proportion of cross-national analyses within these articles that account for non-independence between nations. Third, we run simulations to determine whether common methods of dealing with non-independence in the literature sufficiently reduce false positive rates. Fourth, we reanalyse twelve previous cross-national analyses of economic development and cultural values from our systematic review, incorporating global geographic and linguistic proximity matrices to correctly control for spatial and cultural non-independence.

Results

National-level economic development and cultural values are spatially and culturally non-independent

In order to motivate our research question, it is important to first quantify the degree of spatial and cultural non-independence for economic development and cultural values around the world. If these variables are only weakly non-independent, then the issue might be safe to ignore. However, if they are strongly non-independent, then there is a possibility that non-independence could be confounding cross-national inferences.

To this end, we used Bayesian multilevel models to simultaneously estimate geographic and cultural phylogenetic signal for a range of economic development and cultural values variables that are widely used in the literature. For economic development, we focused on the Human Development Index[45], gross domestic product per capita, gross domestic product per capita growth, and the Gini index of income inequality. For cultural values, we focused on two primary dimensions of cultural values from the World Values Survey, traditional vs. secular values and survival vs. self-expression values[16], as well as cultural tightness[14] and individualism[15].

For all of these variables, we found that a substantial proportion of national-level variation was explained by spatial proximity and/or shared cultural ancestry between nations (Figure 1; Supplementary Table S1). Signal estimates were often markedly strong, with spatial proximity and shared cultural ancestry frequently explaining over half of the national-level variation. For spatial proximity, Bayes Factors indicated strong evidence that the geographic signal estimates differed from zero for all economic development variables and traditional values. However, the evidence was only equivocal for survival values and individualism, and strong evidence was found that the geographic signal estimate for tightness was equal to zero. For shared cultural ancestry, Bayes Factors indicated strong evidence that the cultural phylogenetic signal estimates differed from zero for all economic development and cultural values variables except for gross domestic product per capita growth, for which the evidence was equivocal. These findings emphasise the need to account for spatial and cultural phylogenetic non-independence in cross-national analyses of economic development and cultural values.

Previous cross-national analyses have not sufficiently accounted for non-independence

Given that economic development and cultural values show evidence of geographic and cultural phylogenetic signal, have cross-national analyses sufficiently accounted for this non-independence? To assess this, we systematically searched the published literature for articles that combined the search terms "economic development" or "values" with the search terms "cross-national", "cross-cultural", or "cross-country". We removed articles that did not report original research, were not relevant to economic development or cultural values, or did not report at least one cross-national analysis. We then retained the 100 articles (50 for economic development, 50 for cultural values) with the highest annual rate of citations (Supplementary Table S2). For each of these highly-cited articles, we exhaustively recorded every cross-national analysis reported in the main text (n = 4,308),

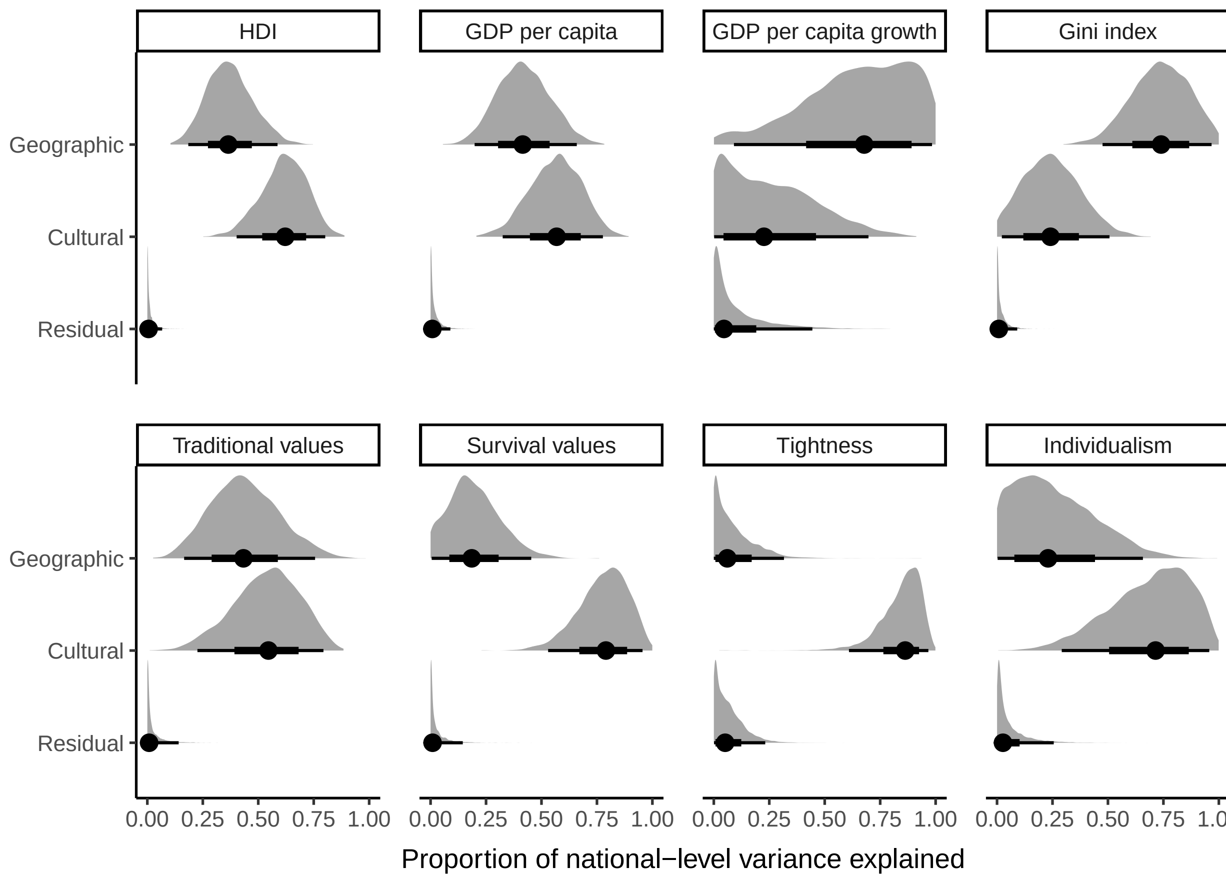


Figure 1. Posterior estimates of geographic and cultural phylogenetic signal for a range of economic development and cultural values variables. Geographic and cultural phylogenetic signal are operationalised as the proportion of national-level variance explained by geographic and linguistic proximity matrices. Grey ridges are full posterior distributions, points are posterior median values, and black lines are 50% and 95% credible intervals. HDI = Human Development Index; GDP = gross domestic product.

identifying in each case whether or not the analysis attempted to control for spatial, cultural, or any other form of non-independence between nations (see Methods for detailed search criteria and coding decisions).

The results of our systematic review show that most published articles containing cross-national analyses make no attempt to account for statistical non-independence. Figure 2a plots the proportion of articles that contain at least one cross-national analysis accounting for non-independence. We find that 42% of economic development articles contain at least one attempt to control for non-independence (95% bootstrap confidence interval [0.30 0.54]), while this proportion decreases to only 8% for cultural values articles (95% bCI [0.02 0.16]). Both kinds of article are most likely to use regional fixed effects (e.g. continent fixed effects) to account for non-independence, but some articles also include controls for spatial distance (e.g. latitude) and shared cultural history (e.g. colony status). These proportions are even lower when focusing on the full sample of 4,308 analyses: only 5% (95% credible interval [0.02 0.13]) of individual economic development analyses and 1% (95% CI [0.00 0.02]) of individual cultural values analyses are estimated to control for non-independence (Supplementary Figure S2).

While our review contains a range of articles from low- and high-impact journals, an anonymous reviewer suggested that our estimates could be biased downwards by analyses published in lower impact outlets with more relaxed standards for issues like non-independence. It is also possible that, since our systematic review goes back as far as 1993, our estimates are being biased downwards by earlier studies, and that controls for non-independence have increased over time with methodological advancements and greater awareness of the issue. To test these possible explanations for our low estimates, we fitted Bayesian logistic regression models to the data from the review, including log journal impact factor and publication year as separate predictors. Interestingly, we found that, contrary to the explanation above, studies from higher impact journals were *less* likely to include at least one control for non-independence than studies from lower impact journals,

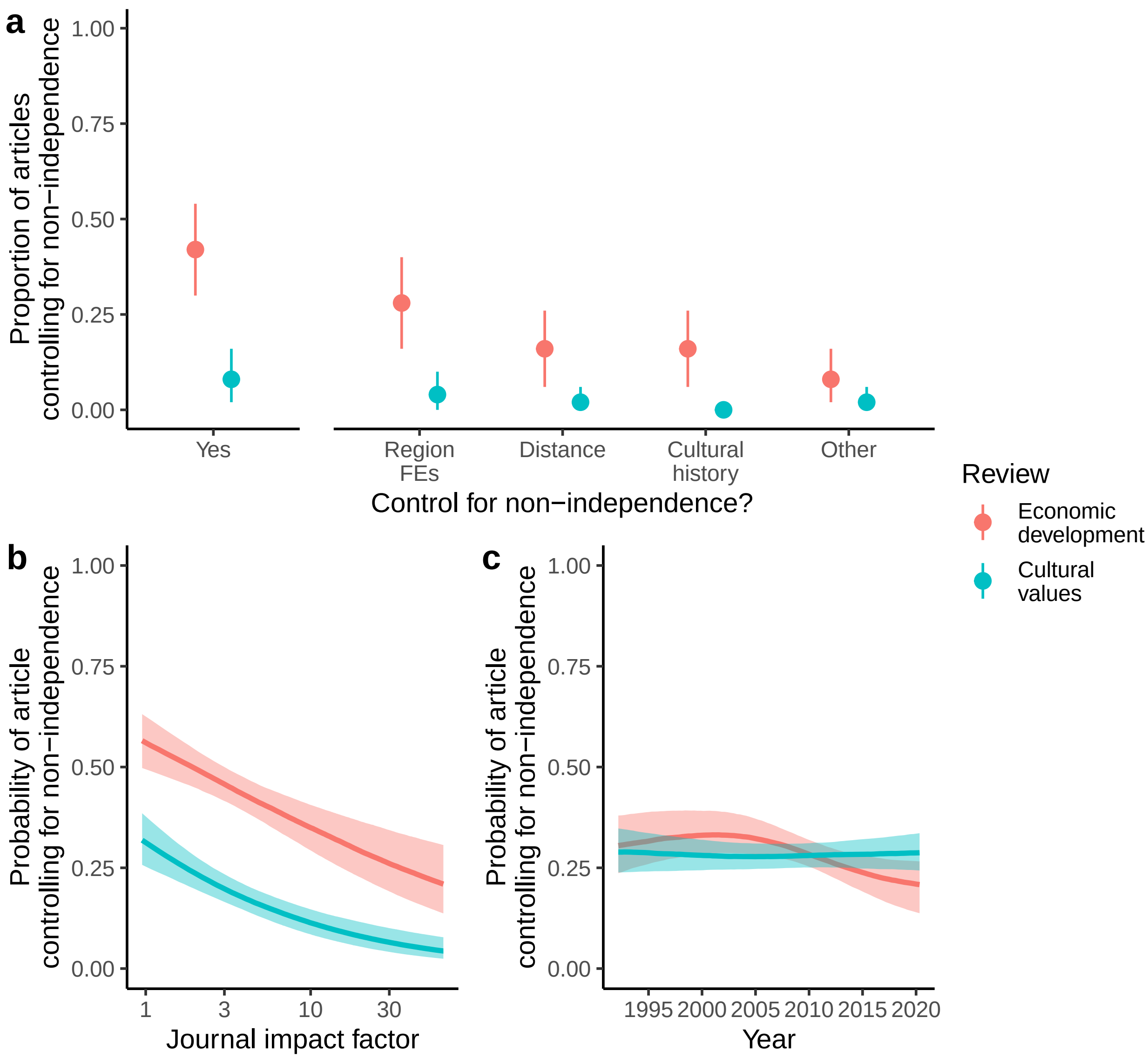


Figure 2. *Results from systematic review of 100 highly-cited cross-national studies of economic development (red) and cultural values (blue).* (a) Proportion of articles containing at least one analysis accounting for non-independence, overall and split by common methods of controlling for non-independence. (b) The association between journal impact factor and the probability that an article contains at least one analysis accounting for non-independence. (c) Estimated trend over time for the probability that an article contains at least one analysis accounting for non-independence. Point ranges represent proportions and 95% bootstrap confidence intervals. Lines and shaded areas are posterior median regression lines and 50% credible intervals from Bayesian multilevel models. Region FEs = region fixed effects.

both for studies of economic development (b = -0.38, 95% CI [-0.87 0.09]) and for studies of cultural values (b = -0.56, 95% CI [-1.10 -0.05]; Figure 2b). Moreover, splines revealed no relationship between publication year and the probability of including at least one control for non-independence, both for studies of economic development (b = -0.13, 95% CI [-1.10 0.82]) and for studies of cultural values (b = -0.02, 95% CI [-1.00 0.96]; Figure 2c).

Common methods of controlling for non-independence are insufficient for reducing false positive rates in non-independent data

Our systematic review revealed that most cross-national analyses in the literature do not control for spatial or cultural phylogenetic non-independence. When they do, they tend to include controls like latitude and regional fixed effects. Do these methods sufficiently account for statistical non-independence?

To compare the efficacy of different methods in the literature, we conducted a simulation study. We simulated national-level datasets (n = 236 nations) with varying degrees of spatial or cultural phylogenetic autocorrelation (i.e. non-independence) for outcome and predictor variables, but with no direct causal relationship between the variables. We then fitted naive regressions without controls to these datasets, as well as regression models with controls for latitude, longitude, and continent fixed effects. Despite not being identified in our systematic review, we also included other methods that are often used in the literature to account for non-independence. Additional spatial controls included the mean of the predictor variable within a surrounding 2000km radius (e.g.[46]) and Conley standard errors[47,48] based on geographic distances between nations (e.g.[46,**Schulz2021?**]). Additional cultural controls included fixed effects for the language families of the majority-spoken languages in each nation (e.g.[49]) and Conley standard errors based on genetic distances between nations (e.g.[46,**Schulz2021?**]). These frequentist approaches attempt to account for non-independence by holding geographic location constant (latitude, longitude), discarding between-region variation and exploiting only local variation (continent fixed effects, mean of surrounding 2000km), or correcting standard errors for autocorrelation post-hoc while leaving model coefficients unchanged (Conley standard errors).

Beyond frequentist approaches, we also fitted Bayesian multilevel regressions that explicitly model spatial and/or cultural phylogenetic non-independence by allowing nations to covary according to geographic and/or linguistic proximity matrices. Geographic proximity between nations is calculated from inverse distances between longitude and latitude coordinates. Linguistic proximity between nations is calculated from a global phylogenetic tree that represents hierarchical relationships of genealogical descent for all languages in the world. For each pair of nations, we calculate inverse phylogenetic distances (i.e. number of branches separating two taxa) between all languages spoken in that nation pair and produce an average "linguistic proximity" score weighted by the percentages of speakers within those nations. To include the resulting geographic and linguistic proximity matrices in our models, we included a Gaussian process[50,51] over latitude and longitude values and/or assumed that nation random intercepts were correlated in proportion to their linguistic proximity[52]. These approaches attempt to account for non-independence by modelling the covariance between nations that is induced by their geographic or linguistic connections.

Figures 3 and 4 plot the estimated false positive rates from our simulation study, split by different methods and different degrees of spatial or cultural phylogenetic autocorrelation (see Supplementary Figures S3 and S4 for full distributions of effect sizes under strong autocorrelation). Across all model types, false positive rates were measured as the proportion of models that estimated a slope with a 95% confidence / credible interval excluding zero (i.e. falsely infering a relationship when none is present). For reference, "weak" autocorrelation in our simulation is comparable to the geographic signal for survival values in Figure 1 (i.e. 20% of the national-level variance is explained by non-independence), while "moderate" and "strong" levels of autocorrelation are comparable to the cultural phylogenetic signal for traditional and survival values, respectively (i.e. 50% and 80% of the national-level variance is explained by non-independence).

Our simulation study revealed that with at least moderate degrees of spatial or cultural phylogenetic autocorrelation for both outcome and predictor variables, naive regression models produce false positive rates above chance levels. This false positive rate increases as the degree of autocorrelation increases. With strong spatial autocorrelation for both outcomes and predictors, false positive rates reach as high as 78%. We find a slightly lower false positive rate under strong cultural phylogenetic autocorrelation, though this false positive rate is still greater than expected by chance (39%).

Common methods in the literature do not reduce these high false positive rates. With strong spatial autocorrelation for both outcome and predictor variables, false positive rates remain above 50% when controlling for latitude, longitude, and language family fixed effects (Figure 3). Applying Conley standard errors also does not reduce false positive rates below 50% under strong spatial autocorrelation, regardless of the distance cutoff. Continent fixed effects are more effective than other frequentist methods, though they continue to produce a false positive rate of 56% under strong spatial autocorrelation. By contrast, Bayesian spatial Gaussian process regression with longitude and latitude values outperforms all other methods. This approach eliminates false positives under moderate spatial autocorrelation, such that the false positive rate is no different from chance, and reduces the false positive rate under strong spatial autocorrelation to 54%. Bayesian models that additionally account for linguistic proximity between nations perform equally well, though models with only linguistic covariance continue to produce false positives.

In our simulation of cultural phylogenetic non-independence, we find that most frequentist methods do not reduce false positive rates (Figure 4). Controls for latitude and longitude, continent fixed effects, and Conley standard errors do little to change false positive rates. Language family fixed effects are slightly more effective than other

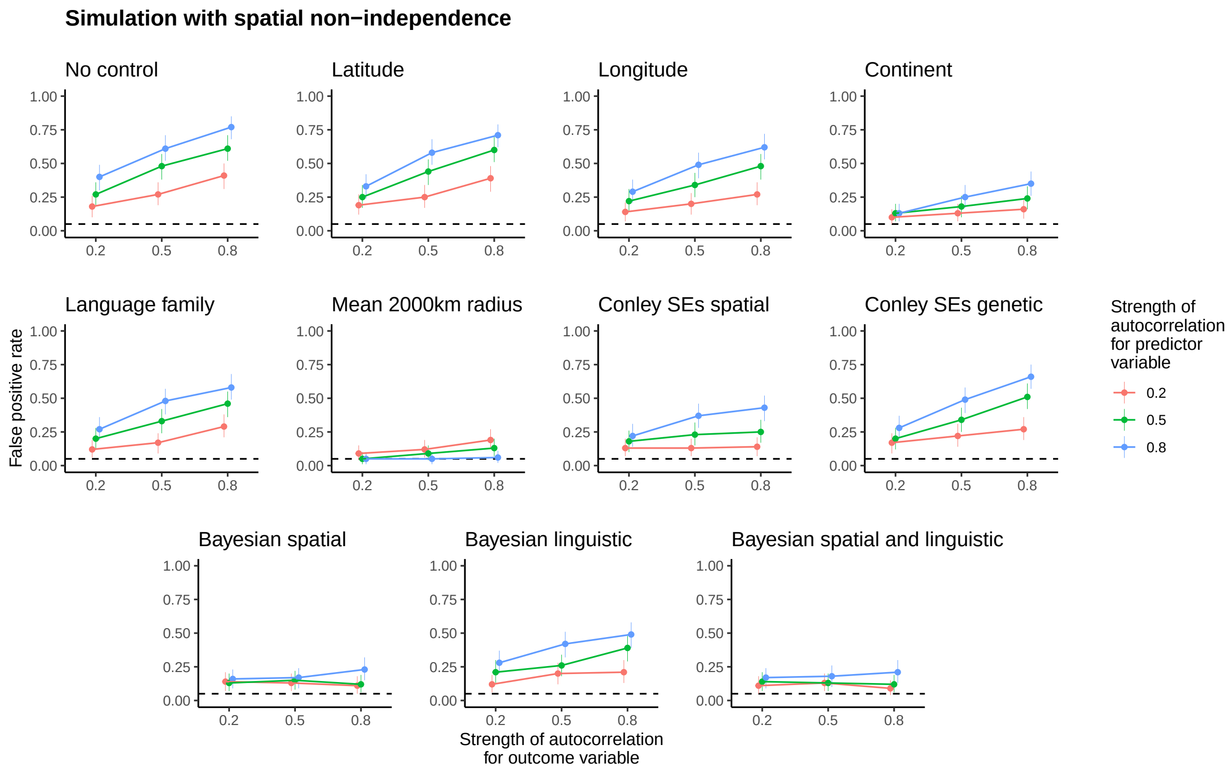


Figure 3. *False positive rates for different methods of controlling for spatial non-independence in our simulation study.* For simulated outcome and predictor variables, we systematically varied the strength of spatial autocorrelation, from weak (0.2) to moderate (0.5) to strong (0.8). We simulated 100 datasets per parameter combination and fitted different models to each dataset. False positive rates were operationalised as the proportion of models that estimated a slope with a 95% confidence / credible interval excluding zero. Point ranges represent proportions and 95% bootstrap confidence intervals, and dashed lines indicate the 5% false positive rate that is expected due to chance. SEs = standard errors.

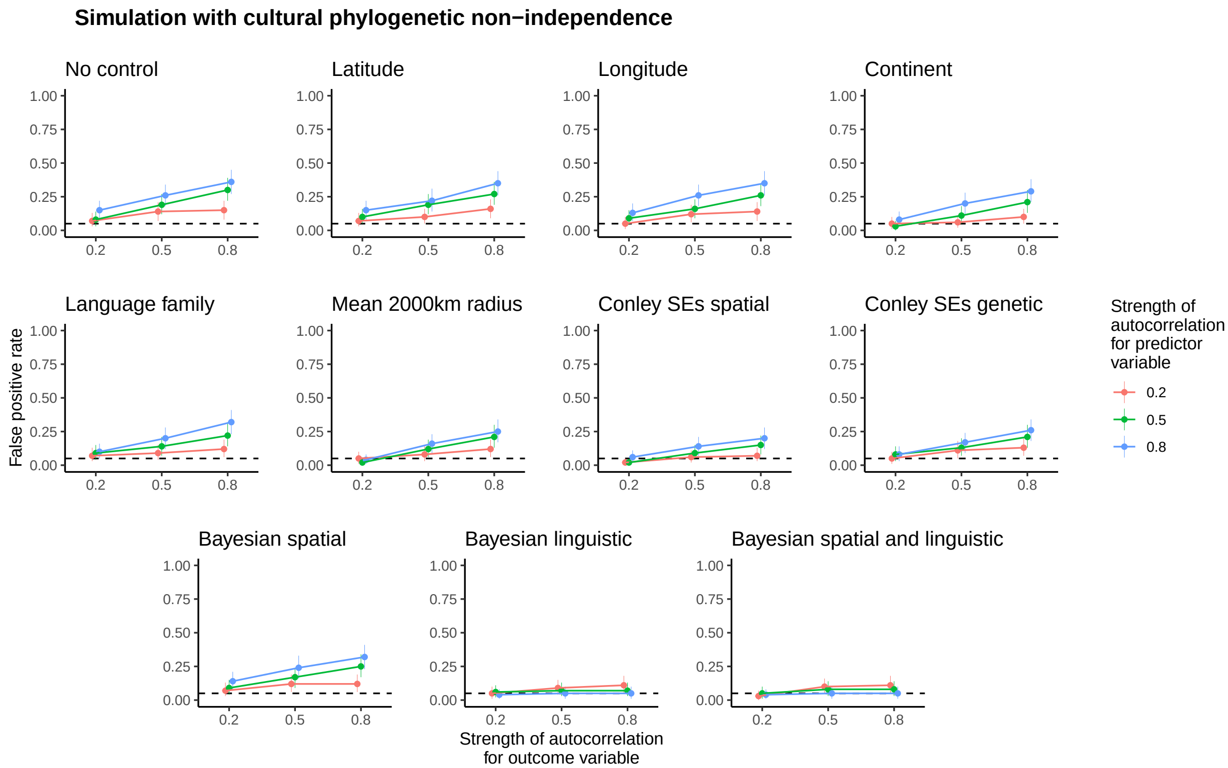


Figure 4. *False positive rates for different methods of controlling for cultural phylogenetic non-independence in our simulation study.* For simulated outcome and predictor variables, we systematically varied the strength of cultural phylogenetic autocorrelation, from weak (0.2) to moderate (0.5) to strong (0.8). We simulated 100 datasets per parameter combination and fitted different models to each dataset. False positive rates were operationalised as the proportion of models that estimated a slope with a 95% confidence / credible interval excluding zero. Point ranges represent proportions and 95% bootstrap confidence intervals, and dashed lines indicate the 5% false positive rate that is expected due to chance. SEs = standard errors.

frequentist methods, though they continue to produce a false positive rate of 39% under strong cultural phylogenetic autocorrelation. By contrast, Bayesian models with random effects covarying according to linguistic proximity completely eliminate false positives across all degrees of cultural phylogenetic autocorrelation. Bayesian models that additionally account for geographic proximity between nations perform equally well, though models with only a spatial Gaussian process continue to produce false positives.

Key findings in the literature are not robust to reanalysis with more rigorous methods

Our systematic review and simulation study have shown that controls for non-independence are rare in cross-national studies of economic development and cultural values, and when studies do attempt to control for non-independence, the methods typically used are likely to continue to produce false positives. This raises the worrying possibility that the cross-national literature in economics and psychology is populated with spurious relationships.

To determine how widespread this issue of spurious cross-national relationships might be, we reanalysed a subset of twelve previous cross-national analyses from our systematic review, sufficiently controlling for spatial and cultural phylogenetic non-independence using global geographic and linguistic proximity matrices. We subsampled six analyses from our economic development review[53–58] and six from our cultural values review[13,14,16,59–61]. Our choice of analyses was constrained by data availability and whether we were able to initially replicate the original finding. We pre-registered our subsample of analyses before running any control models (https://osf.io/uywx8/). We controlled for non-independence by including (1) a Gaussian process allowing nation random intercepts to covary according to a geographic proximity matrix, and/or (2) nation random intercepts that covaried according to a linguistic proximity matrix (see Supplementary Methods for full models).

Figure 5 visualises the results of our reanalysis. Cross-national correlation effect sizes tended to reduce when controlling for statistical non-independence between nations, sometimes by as much as half of the original effect size. Overall, after controlling for non-independence, six out of twelve cross-national associations had 95% credible intervals that included zero. For the economic development analyses, four out of six cross-national relationships had 95% credible intervals including zero when controlling for spatial non-independence. For the cultural values analyses, two out of six cross-national relationships had 95% credible intervals including zero when controlling for cultural phylogenetic non-independence. Supplementary Figure S5 shows these cross-national correlations plotted against the raw data.

To understand why some cross-national correlations were attenuated by controls for non-independence while others were robust, we further explored our fitted models for evidence of spatial and cultural autocorrelation. For each outcome variable, our Gaussian process models provided varying estimates of how quickly spatial autocorrelation declined with distance (Supplementary Figure S6). For example, in Skidmore and Toya[58] gross domestic product growth was only moderately spatially autocorrelated at 1,000 km distance (posterior median spatial autocorrelation at 1,000 km = 0.42, 95% CI [0.07 0.90]), whereas in Inglehart and Baker[16] traditional values were strongly spatially autocorrelated at the same distance (posterior median spatial autocorrelation at 1,000 km = 0.96, 95% CI [0.81 0.99]). We also found varying estimates of cultural phylogenetic signal (Supplementary Figure S7), with some outcome variables expressing low signal (e.g. tightness[14]; posterior median = 0.15, 95% CI [0.00 0.85]) and others expressing high signal (e.g. female labour force participation[60]; posterior median = 0.89, 95% CI [0.63 0.98]). Across all analyses, we found that stronger estimates of spatial autocorrelation or cultural phylogenetic signal resulted in a more pronounced reduction in the effect size when controlling for non-independence between nations (Figure 6).

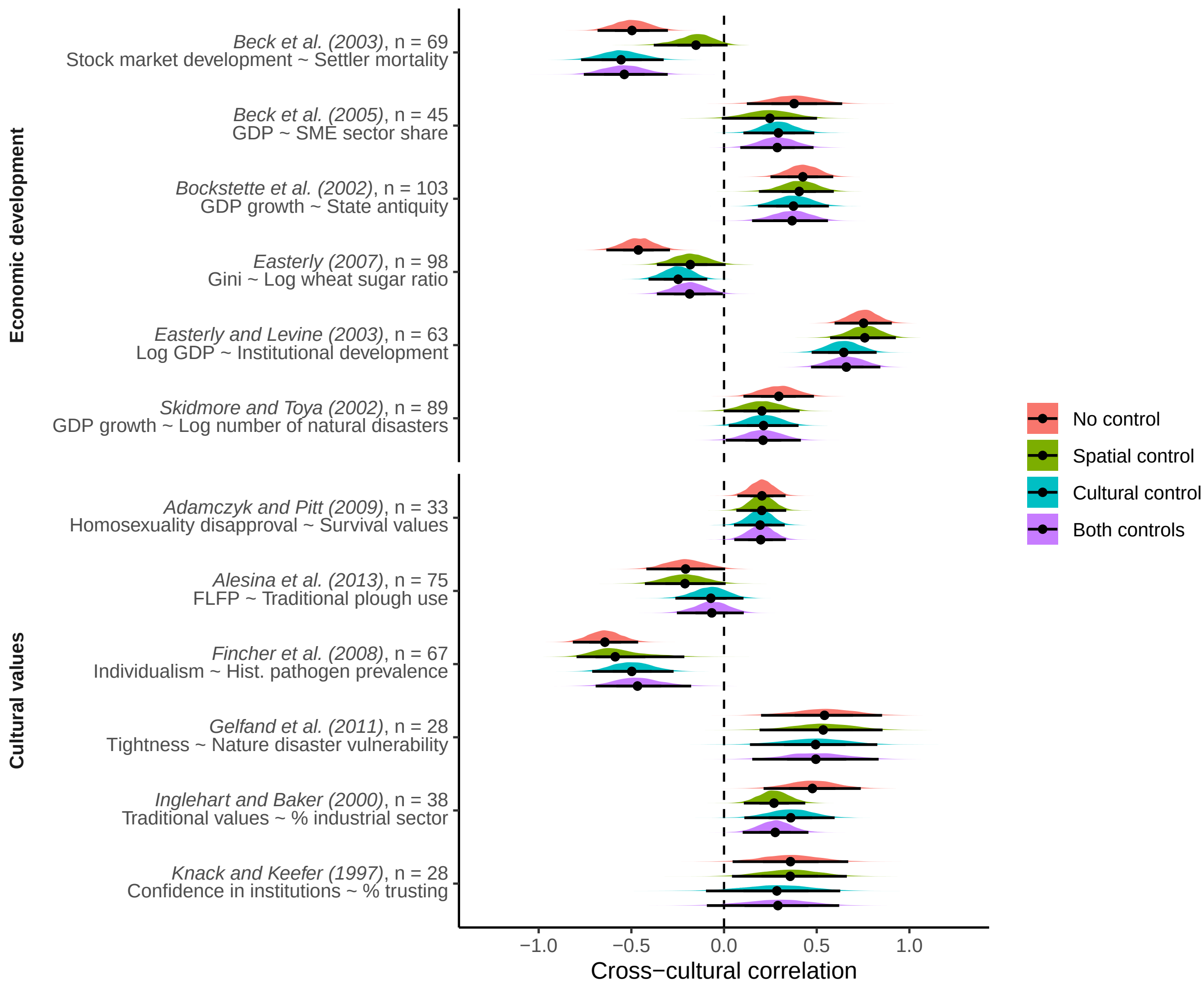


Figure 5. *Posterior correlations from our reanalysis of twelve previous cross-national analyses.* For each previous cross-national relationship, we plot the posterior slopes from a naive regression (red), a regression controlling for spatial non-independence (green), a regression controlling for cultural phylogenetic non-independence (blue), and a regression controlling for both spatial and cultural phylogenetic non-independence simultaneously (purple). All outcome and predictor variables are standardised. Most analyses are simple bivariate cross-national correlations, but Gelfand et al. (2011) is a partial correlation controlling for log gross national income and Adamczyk and Pitt (2009) is a multilevel model including several covariates. Points and black lines represent posterior medians and 95% credible intervals. GDP = gross domestic product. FLFP = female labour force participation.

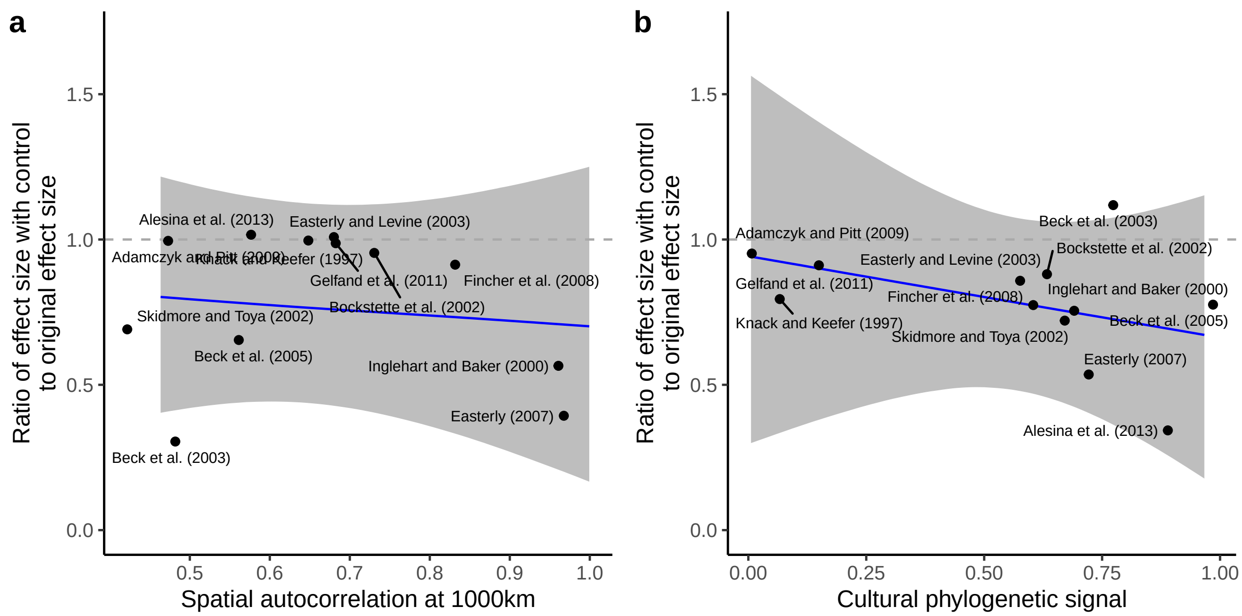


Figure 6. The estimated degree of spatial and cultural phylogenetic non-independence predicts reductions in effect size in our reanalysis. (a) Higher estimated degrees of spatial autocorrelation at 1,000 km distance predict more pronounced reductions in effect sizes when controlling for non-independence. (b) Higher estimated levels of cultural phylogenetic signal predict more pronounced reductions in effect sizes when controlling for non-independence. In both panels, the y-axis represents the ratio of the effect size when controlling for spatial and cultural non-independence to the original effect size (from naive regression model), and the x-axis represents posterior median model estimates. Regression lines are plotted with 95% confidence intervals.

Discussion

In a systematic literature review and simulation, we found that cross-national studies in economics and psychology rarely account for non-independence between nations, and, when they do, the methods they use continue to produce false positives. In a reanalysis of twelve cross-national correlations, we further showed that neglecting to account for non-independence has resulted in spurious relationships in the published literature, with half of the correlations failing to replicate when controlling for spatial or cultural non-independence with more rigorous methods. These findings suggest that cross-national analyses in economics and psychology should be interpreted with caution until non-independence is sufficiently accounted for.

Our initial analyses add to and clarify existing evidence regarding the non-independence of economic and cultural variation among nations. One previous study suggested that geographic proximity is more important than deep cultural ancestry in explaining the distribution of human development across Eurasian nations, though the authors noted that their small sample of 44 nations and regional focus limited their statistical power[62]. Our global sample of nations revealed strong cultural phylogenetic signal, as well as geographic signal, for the Human Development Index. Another previous study found that similarities in the cultural values of nations are predicted by linguistic, but not geographic, distances between those nations[6]. We find this same result for survival vs. self-expression values, but for traditional vs. secular values we find that both linguistic and geographic proximity are important independent predictors of global variation. These findings emphasise the need to account for both spatial and cultural phylogenetic non-independence in cross-national studies of economic development and cultural values.

Crucially, our systematic literature review and simulation study revealed that the most commonly used controls for non-independence do not sufficiently deal with the issue. In our simulations, controlling for either latitude or longitude did not reduce false positive rates. This result calls into question controls like distance to the equator to account for non-independence in cross-national regression models, though these controls may still be suitable to account for regional or latitudinal variation in ecology, which we did not simulate. High false positive rates persisted with Conley standard errors, which have previously been critiqued for being overly sensitive to arbitrary distance cutoffs[40]. The simulation also confirmed the assertion that fixed effects for spatial or cultural groupings (e.g. continent or language family fixed effects) are insufficient because non-independence still remains within groupings[41]. This logic further applies to analyses that control for non-independence by separately analysing different regions (e.g.[63]). Across all model types in our simulation, the only methods that sufficiently reduced the false positive rate were the Bayesian multilevel regressions that explicitly modelled spatial and cultural phylogenetic autocorrelation, though we did not include other possible controls for non-independence, such as conditional autoregressive models[31] or generalised additive models[64].

Ours is not the first review to show that studies are misapplying statistical methods in ways that inflate false positive rates. For example, other literature reviews have shown that studies in the social sciences tend to use small samples of participants[65], treat ordinal data as metric[66], incorrectly handle missing values[67], and ignore best practices in meta-analyses[68]. Why do cross-national studies also rarely account for non-independence? At the institutional level, one possibility is that such practices are incentivised *because* they generate statistically significant relationships, which increase the probability that a study is published[65]. At the individual level, another possibility is that researchers outside of anthropology and ecology are simply not aware of the problem, or believe that the problem does not apply to analyses of nations. Even if researchers appreciate the problem, they might not know of suitable controls or perceive the methods to be too complex.

These institutional- and individual-level barriers can be combatted. First, cross-national replication studies like ours and others[40–44], combined with the methodological reviews included in Registered Reports[69], might change incentive structures and encourage researchers to analyse the world's nations with more rigorous methods. Second, since the issue of non-independence is fundamentally an issue of causal inference (Supplementary Figure S1), more explicit descriptions of causal models could promote controls for non-independence in cross-national research. In our review, economists studying economic development dealt with national-level non-independence more than psychologists studying cultural values, likely because economics studies tend to be lengthy statistical exercises that systematically incorporate or exclude numerous variables in an attempt to infer causation. Third, the recent widespread accessibility of open source statistical software, such as the programming language Stan[70] and the R package *brms*[71], should promote the use of more rigorous methods to control for non-independence. Using *brms*, for example, Bayesian Gaussian process regression is straightforward to conduct, requiring only longitude and latitude values for nations.

Until such changes are implemented and sufficient controls for non-independence are the norm, existing cross-national correlations should be interpreted with caution. In our reanalyses, we found that four out of six cross-national correlations with economic development variables had 95% credible intervals that included zero when controlling for spatial non-independence. Three of these analyses were tests of "persistence" hypotheses, studying the effects of historical and environmental conditions — settler mortality[53], wheat-sugar suitability[57], and natural disaster frequency[58] — on modern developmental outcomes. A recent reanalysis has also called into question various studies of this ilk[40]. We also found that two out of six cross-national correlations with cultural values variables had 95% credible intervals that included zero when controlling for cultural phylogenetic non-independence.

We do not wish to dissuade researchers from conducting cross-national studies. On the contrary, such work promises to deepen understanding of our world, including the causes and consequences of economic development and cultural values. Moreover, cross-national studies allow social scientists to broaden their scope of study beyond Western populations[7], providing the representative samples necessary to test evolutionary and socio-ecological theories of human behaviour[8,72]. But in order to minimise spurious relationships in global datasets, we urge researchers to control for spatial and cultural phylogenetic non-independence when reporting cross-national correlations. Nations are not independent, and our statistical models must reflect this.

Methods

Geographic and cultural phylogenetic signal

To estimate the degree of spatial and cultural phylogenetic non-independence in economic development and cultural values, we calculated geographic and cultural phylogenetic signal for global measures of development and values. Our measure of economic development was the Human Development Index[45]. We retrieved a longitudinal dataset capturing human development for 189 nations since 1990 (n = 1,512; https://hdr.undp.org/en/content/download-data). Our measures of cultural values were traditional vs. secular values and survival vs. self-expression values from the World Values Survey[16]. We downloaded the full Integrated Values Survey, which included all waves from the World Values Survey and the European Values Survey, and computed the two dimensions of cultural values following procedures from previous research[16]. This longitudinal dataset captures values and attitudes for 116 nations since 1981 (n = 645,249; https://www.worldvaluessurvey.org/WVSEVStrend.jsp).

To calculate geographic and cultural phylogenetic signal, we created two proximity matrices for 269 of the world's nations: a geographic proximity matrix and a linguistic proximity matrix. Geographic proximity was converted from logged geodesic distances between nation capital cities. Linguistic proximity was calculated as the cultural proximity between all languages spoken within nations, weighted by speaker percentages (see Supplementary Methods). We included these matrices in Bayesian multilevel models, allowing nation random intercepts to covary according to both geographic and linguistic proximity simultaneously. These models were fitted with the R package *brms*[71] and converged normally ($\hat{R} < 1.1$). Estimates of geographic and cultural phylogenetic signal were computed as the proportion of national-level variance in these models explained by geographic and linguistic proximity matrices.

Systematic literature review

We exported two searches from Web of Science (https://www.webofknowledge.com/) on 27th September 2021, restricting our searches to articles published between 1900 and 2018. The first search was for the terms "economic development" AND ("cross-national" OR "cross-cultural" OR "cross-country"), which returned 965 articles. The second search was for the terms "values" AND ("cross-national" OR "cross-cultural" OR "cross-country"), which returned 6806 articles. Once exported, we ordered the articles by descending number of citations per year since initial publication, using citation counts reported by Web of Science.

We then systematically coded each article, in order, for inclusion in our review. Articles were only included if: (1) they were judged to be relevant to economic development or cultural values; (2) they were an original empirical research article; and (3) they contained at least one analysis with national-level outcome or predictor variables. We stopped when we had included 50 articles for the economic development review and 50 articles for the cultural values review.

Within each included article, we exhaustively coded every individual cross-national analysis reported in the main text. We coded mainly correlation or regression analyses, and explicitly excluded meta-analyses, factor analyses, measurement invariance analyses, multidimensional scaling analyses, hierarchical clustering analyses, multiverse analyses, and scale development / validation analyses. We also excluded analyses that compared only two, three, four, five, or six nations. For each included analysis, we recorded the year, outcome variable, all predictor variables, test statistic, p-value, number of nations, number of data points, model type, if the data were available, and whether and how the analysis attempted to control for non-independence.

We coded common attempts to control for non-independence between nations. These included: (1) any higher-level control variables for spatial regional groupings (e.g. continent fixed effects); (2) any geographic distance control variables (e.g. distance between capital cities, distance from equator, latitude); (3) any control variables capturing shared cultural history (e.g. former colony, legal origin fixed effects, linguistic history, cultural influence); and (4) any other control variables, tests, or approaches that were deemed as attempts to control for non-independence (e.g. eigenvector filtering[73], controls for trade-weightings between nations, cross-sectional dependence tests[74], separate analyses for subsets of nations). These were coded by the first author.

Once we had compiled our review database, we calculated the proportion of articles attempting to control for non-independence at least once. We also calculated the proportion of articles employing the different types of control listed above at least once: regional fixed effects, distance, shared cultural history, or other. For these proportions, we calculated 95% boostrap confidence intervals with 1,000 bootstrap iterations.

For individual analyses, we dealt with the nested nature of the data (analyses nested within articles) by fitting Bayesian multilevel logistic regression models with review type (economic development vs. cultural values) as the sole fixed effect and random intercepts for articles. We fitted these models separately for overall attempts to control for non-independence and split by method type. We report the adjusted proportions with 95% credible intervals. To test for a trend over time, we also fitted a Bayesian multilevel logistic regression with a multigroup spline for year of publication and random intercepts for articles. Bayesian models were fitted with the *brms* R package[71]. Our priors were informed by prior predictive checks, and all models converged normally ($\hat{R} < 1.1$).

Simulations

We simulated data for 236 nations with varying degrees of spatial or cultural phylogenetic signal for outcome y and predictor x using the following generative model:

$$y = \alpha_y + \epsilon_y$$
$$x = \alpha_x + \epsilon_x$$
$$\alpha_y \sim \mathcal{N}(0, \sqrt{\lambda} \cdot \Sigma)$$
$$\alpha_x \sim \mathcal{N}(0, \sqrt{\rho} \cdot \Sigma)$$
$$\epsilon_y \sim \mathcal{N}(0, \sqrt{1-\lambda})$$
$$\epsilon_x \sim \mathcal{N}(0, \sqrt{1-\rho})$$

where Σ is a correlation matrix proportional to either geographic or linguistic proximities between nations, and λ and ρ are autocorrelation parameters that represent the expected spatial or cultural phylogenetic signal for outcome and predictor variables, respectively. Importantly, in this simulation, we know that there is no direct causal relationship between y and x because we simulate the variables independently. Instead, any relationship between the two variables is merely the result of autocorrelation.

We set the autocorrelation parameters to either 0.2 (weak), 0.5 (moderate), or 0.8 (strong). We simulated 100 datasets for each parameter combination, resulting in 900 datasets. Each dataset had 236 rows representing different nations, with the following associated data for each nation: latitude, longitude, continent (Africa, Asia, Europe, North America, Oceania, or South America), and language family of the nation's majority spoken language (Afro-Asiatic, Atlantic-Congo, Austroasiatic, Austronesian, Eskimo-Aleut, Indo-European, Japonic, Kartvelian, Koreanic, Mande, Mongolic-Khitan, Nilotic, Nuclear Trans New Guinea, Sino-Tibetan, Tai-Kadai, Tupian, Turkic, or Uralic).

With the resulting simulated datasets, we standardised outcome and predictor variables and fitted eleven different models: (1) naive regression without controls, (2) regression with latitude control, (3) regression with longitude control, (4) regression with continent fixed effects, (5) regression with language family fixed effects, (6) regression employing Conley standard errors with 100 km cutoff, (7) regression employing Conley standard errors with 1,000 km cutoff, (8) regression employing Conley standard errors with 10,000 km cutoff, (9) Bayesian regression including a Gaussian process over latitudes and longitudes, (10) Bayesian regression including random intercepts covarying according to linguistic proximity, and (11) Bayesian regression including both a Gaussian process over latitudes and longitudes and random intercepts covarying according to linguistic proximity.

Models employing Conley standard errors required latitude and longitude values, and the cutoffs implied the distance beyond which autocorrelation is assumed to be zero. These models were fitted using the *conleyreg* R package[75]. Bayesian models were fitted using the *brms* R package[71]. Our choice of priors was based on prior predictive simulation. All models converged normally ($\hat{R} < 1.1$). Across all model types and parameter combinations, we calculated the false positive rate as the proportion of models that estimated slopes with a 95% confidence / credible interval excluding zero. We calculated 95% bootstrap confidence intervals for these false positive rates with 1,000 bootstrap iterations.

Reanalyses

We searched the individual analyses from our systematic review for cross-national correlations with available data. We included only analyses for which we were able to replicate the original result (i.e. find a cross-national correlation with the same sign and roughly the same effect size). We restricted our search to one analysis per article, and aimed for an even number of analyses for both economic development and cultural values studies. We also ensured that at least one analysis was a multilevel model, with multiple observations per nation.

The twelve analyses that we settled on[13,14,16,53–61] were mostly bivariate cross-national correlations, except for two. One analysis[14] additionally controlled for log gross national income, and another analysis[58] is a multilevel model including random intercepts for nations and several individual-level and national-level covariates (see Model 5 in original paper). Before running any additional models, we pre-registered these twelve analyses on the Open Science Framework on 25th January 2022 (https://osf.io/uywx8/).

For each individual analysis, we ran four models: (1) a naive regression replicating the original finding, (2) a regression including a Gaussian process allowing nation random intercepts to covary according to a geographic proximity matrix from latitude and longitude values, (3) a regression including nation random intercepts that covaried according to a linguistic proximity matrix, and (4) a regression including both a geographic Gaussian process and nation random intercepts with linguistic covariance. See Supplementary Methods for full models.

We fitted these models using the *brms* R package[71]. Our choice of priors was based on prior predictive simulation. All models converged normally ($\hat{R} < 1.1$), though for some models we resorted to using approximate Gaussian processes[76] to reach convergence.

Reproducibility

All data and code are accesible at our Open Science Framework repository (https://osf.io/uywx8/). We used the *targets* R package[77] to create a reproducible data analysis pipeline and the *papaja* R package[78] to reproducibly generate the manuscript.

References

1. Caselli, F. Accounting for cross-country income differences. in (eds. Aghion, P. & Durlauf, S. N.) vol. 1 679–741 (Elsevier, 2005).
2. Austin, K. F. & McKinney, L. A. Disease, war, hunger, and deprivation: A cross-national investigation of the determinants of life expectancy in less-developed and sub-Saharan African nations. *Sociological Perspectives* **55**, 421–447 (2012).
3. Rai, D., Zitko, P., Jones, K., Lynch, J. & Araya, R. Country- and individual-level socioeconomic determinants of depression: Multilevel cross-national comparison. *British Journal of Psychiatry* **202**, 195–203 (2013).
4. Diener, E., Diener, M. & Diener, C. Factors predicting the subjective well-being of nations. in *Culture and well-being: The collected works of Ed Diener* (ed. Diener, E.) 43–70 (Springer Netherlands, 2009). doi:10.1007/978-90-481-2352-0_3.
5. Kirkcaldy, B., Furnham, A. & Siefen, G. The relationship between health efficacy, educational attainment, and well-being among 30 nations. *European Psychologist* **9**, 107–119 (2004).
6. White, C. J. M., Muthukrishna, M. & Norenzayan, A. Cultural similarity among coreligionists within and between countries. *Proceedings of the National Academy of Sciences* **118**, e2109650118 (2021).
7. Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behavioral and Brain Sciences* **33**, 61–83 (2010).
8. Pollet, T. V., Tybur, J. M., Frankenhuis, W. E. & Rickard, I. J. What can cross-cultural correlations teach us about human nature? *Human Nature* **25**, 410–429 (2014).
9. Benhabib, J. & Spiegel, M. M. The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics* **34**, 143–173 (1994).

10. Comin, D., Easterly, W. & Gong, E. Was the wealth of nations determined in 1000 BC? *American Economic Journal: Macroeconomics* **2**, 65–97 (2010).

11. La Porta, R., Lopez-De-Silanes, F., Shleifer, A. & Vishny, R. W. Legal determinants of external finance. *The Journal of Finance* **52**, 1131–1150 (1997).

12. Sachs, J. D. & Warner, A. M. The curse of natural resources. *European Economic Review* **45**, 827–838 (2001).

13. Fincher, C. L., Thornhill, R., Murray, D. R. & Schaller, M. Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. *Proceedings of the Royal Society B: Biological Sciences* **275**, 1279–1285 (2008).

14. Gelfand, M. J. *et al.* Differences between tight and loose cultures: A 33-nation study. *Science* **332**, 1100–1104 (2011).

15. Hofstede, G. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations.* (Sage Publications, 2001).

16. Inglehart, R. & Baker, W. E. Modernization, cultural change, and the persistence of traditional values. *American Sociological Review* **65**, 19–51 (2000).

17. Awad, E. *et al.* The moral machine experiment. *Nature* **563**, 59–64 (2018).

18. Rhoads, S. A., Gunter, D., Ryan, R. M. & Marsh, A. A. Global variation in subjective well-being predicts seven forms of altruism. *Psychological Science* **32**, 1247–1261 (2021).

19. Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P. & Henrich, J. The Church, intensive kinship, and global psychological variation. *Science* **366**, eaau5141 (2019).

20. Thomson, R. *et al.* Relational mobility predicts social behaviors in 39 countries and is tied to historical farming and threat. *Proceedings of the National Academy of Sciences* **115**, 7521–7526 (2018).

21. Kissling, W. D. & Carl, G. Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography* **17**, 59–71 (2008).

22. Hewlett, B. S., DeSilvestri, A. & Guglielmino, C. R. Semes and genes in africa. *Current Anthropology* **43**, 313–321 (2002).

23. Nunn, N. & Puga, D. Ruggedness: The blessing of bad geography in Africa. *The Review of Economics and Statistics* **94**, 20–36 (2012).

24. Tobler, W. R. A computer movie simulating urban growth in the Detroit region. *Economic Geography* **46**, 234–240 (1970).

25. Tylor, E. B. On a method of investigating the development of institutions; applied to laws of marriage and descent. *The Journal of the Anthropological Institute of Great Britain and Ireland* **18**, 245–272 (1889).

26. Naroll, R. Two solutions to Galton's Problem. *Philosophy of Science* **28**, 15–39 (1961).

27. Naroll, R. Galton's Problem: The logic of cross-cultural analysis. *Social Research* **32**, 428–451 (1965).

28. Jarque, C. M. & Bera, A. K. A test for normality of observations and regression residuals. *International Statistical Review / Revue Internationale de Statistique* **55**, 163–172 (1987).

29. Legendre, P. Spatial autocorrelation: Trouble or new paradigm? *Ecology* **74**, 1659–1673 (1993).

30. Roberts, J., Seán AND Winters. Linguistic diversity and traffic accidents: Lessons from statistical studies of cultural traits. *PLOS ONE* **8**, 1–13 (2013).

31. Lichstein, J. W., Simons, T. R., Shriner, S. A. & Franzreb, K. E. Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs* **72**, 445–463 (2002).

32. Loftin, C. & Ward, S. K. A spatial autocorrelation model of the effects of population density on fertility. *American Sociological Review* **48**, 121–128 (1983).

33. Ferguson, J. Anthropology and its evil twin. in (eds. Cooper, F. & Packard, R.) 150–175 (University of California Press Berkeley, 1997).

34. Murdock, G. P. & White, D. R. Standard cross-cultural sample. *Ethnology* **8**, 329–369 (1969).

35. Dow, M. M. & Eff, E. A. Global, regional, and local network autocorrelation in the standard cross-cultural sample. *Cross-Cultural Research* **42**, 148–171 (2008).

36. Eff, E. A. Does Mr. Galton still have a problem? Autocorrelation in the standard cross-cultural sample. *World Cultures* **15**, 153–170 (2004).

37. Symonds, M. R. E. & Blomberg, S. P. A primer on phylogenetic generalised least squares. in *Modern phylogenetic comparative methods and their application in evolutionary biology: Concepts and practice* (ed. Garamszegi, L. Z.) 105–130 (Springer Berlin Heidelberg, 2014). doi:10.1007/978-3-662-43550-2_5.

38. Watts, J., Sheehan, O., Bulbulia, J., Gray, R. D. & Atkinson, Q. D. Christianity spread faster in small, politically structured societies. *Nature Human Behaviour* **2**, 559–564 (2018).

39. Atkinson, Q. D., Coomber, T., Passmore, S., Greenhill, S. J. & Kushnick, G. Cultural and environmental predictors of pre-european deforestation on pacific islands. *PLOS ONE* **11**, 1–15 (2016).

40. Kelly, M. Understanding persistence. *CEPR Discussion Paper No. DP15246* (2020).

41. Bromham, L., Hua, X., Cardillo, M., Schneemann, H. & Greenhill, S. J. Parasites and politics: Why cross-cultural studies must control for relatedness, proximity and covariation. *Royal Society Open Science* **5**, 181100 (2018).

42. Bromham, L., Skeels, A., Schneemann, H., Dinnage, R. & Hua, X. There is little evidence that spicy food in hot countries is an adaptation to reducing infection risk. *Nature Human Behaviour* **5**, 878–891 (2021).

43. Currie, T. E. & Mace, R. Analyses do not support the parasite-stress theory of human sociality. *Behavioral and Brain Sciences* **35**, 83–85 (2012).

44. Passmore, S. & Watts, J. WEIRD people and the Western Church: Who made whom? *Religion, Brain & Behavior* 1–58 (2022) doi:10.1080/2153599X.2021.1991459.

45. *Human development report.* http://hdr.undp.org/en/composite/HDI.

46. Schulz, J. F. Kin networks and institutional development. *The Economic Journal* **132**, 2578–2613 (2022).

47. Conley, T. G. GMM estimation with cross sectional dependence. *Journal of Econometrics* **92**, 1–45 (1999).

48. Conley, T. G. Spatial econometrics. in *Microeconometrics* (eds. Durlauf, S. N. & Blume, L. E.) 303–313 (Palgrave Macmillan, 2010).

49. Currie, T. E. & Mace, R. Political complexity predicts the spread of ethnolinguistic groups. *Proceedings of the National Academy of Sciences* **106**, 7339–7344 (2009).

50. McElreath, R. *Statistical rethinking: A Bayesian course with examples in R and Stan.* (CRC Press, 2020).

51. Neal, R. M. Regression and classification using Gaussian process priors. in *Bayesian statistics* (eds. Bernardo, J. M., Berger, J. O., Dawid, A. P. & Smith, A. F. M.) vol. 6 475–501 (Oxford University Press).

52. Villemereuil, P. de & Nakagawa, S. General quantitative genetic methods for comparative biology. in *Modern phylogenetic comparative methods and their application in evolutionary biology: Concepts and practice* (ed. Garamszegi, L. Z.) 287–303 (Springer Berlin Heidelberg, 2014). doi:10.1007/978-3-662-43550-2_11.

53. Beck, T., Demirgäç-Kunt, A. & Levine, R. Law, endowments, and finance. *Journal of Financial Economics* **70**, 137–181 (2003).

54. Beck, T., Demirgäç-Kunt, A. & Levine, R. SMEs, growth, and poverty: Cross-country evidence. *Journal of economic growth* **10**, 199–229 (2005).

55. Bockstette, V., Chanda, A. & Putterman, L. States and markets: The advantage of an early start. *Journal of Economic growth* **7**, 347–369 (2002).

56. Easterly, W. & Levine, R. Tropics, germs, and crops: How endowments influence economic development. *Journal of Monetary Economics* **50**, 3–39 (2003).

57. Easterly, W. Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics* **84**, 755–776 (2007).

58. Skidmore, M. & Toya, H. Do natural disasters promote long-run growth? *Economic Inquiry* **40**, 664–687 (2002).

59. Adamczyk, A. & Pitt, C. Shaping attitudes about homosexuality: The role of religion and cultural context. *Social Science Research* **38**, 338–351 (2009).

60. Alesina, A., Giuliano, P. & Nunn, N. On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics* **128**, 469–530 (2013).

61. Knack, S. & Keefer, P. Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics* **112**, 1251–1288 (1997).

62. Sookias, R. B., Passmore, S. & Atkinson, Q. D. Deep cultural ancestry and human development indicators across nation states. *Royal Society Open Science* **5**, 171411 (2018).

63. Smith, M. D., Rabbitt, M. P. & Coleman- Jensen, A. Who are the world's food insecure? New evidence from the Food and Agriculture Organization's Food Insecurity Experience Scale. *World Development* **93**, 402–412 (2017).

64. Hastie, T. J. & Tibshirani, R. J. *Generalized additive models.* (Routledge, 2017).

65. Smaldino, P. E. & McElreath, R. The natural selection of bad science. *Royal Society Open Science* **3**, 160384 (2016).

66. Liddell, T. M. & Kruschke, J. K. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology* **79**, 328–348 (2018).

67. Nicholson, J. S., Deboeck, P. R. & Howard, W. Attrition in developmental psychology: A review of modern missing data reporting and practices. *International Journal of Behavioral Development* **41**, 143–153 (2017).

68. Tipton, E., Pustejovsky, J. E. & Ahmadi, H. Current practices in meta-regression in psychology, education, and medicine. *Research Synthesis Methods* **10**, 180–194 (2019).

69. Chambers, C. D. & Tzavella, L. The past, present and future of Registered Reports. *Nature Human Behaviour* **6**, 29–42 (2021).

70. Stan Development Team. RStan: The R interface to Stan. (2020).

71. Bürkner, P.-C. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* **80**, 1–28 (2017).

72. Oishi, S. Socioecological psychology. *Annual Review of Psychology* **65**, 581–609 (2014).

73. Tiefelsdorf, M. & Griffith, D. A. Semiparametric filtering of spatial autocorrelation: The eigenvector approach. *Environment and Planning A: Economy and Space* **39**, 1193–1221 (2007).

74. De Hoyos, R. E. & Sarafidis, V. Testing for cross-sectional dependence in panel-data models. *The Stata Journal* **6**, 482–496 (2006).

75. Düben, C. *conleyreg: Estimations using Conley standard errors*. (2021).

76. Solin, A. & Särkkä, S. Hilbert space methods for reduced-rank Gaussian process regression. *Statistics and Computing* **30**, 419–446 (2020).

77. Landau, W. M. The targets R package: A dynamic Make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. *Journal of Open Source Software* **6**, 2959 (2021).

78. Aust, F. & Barth, M. *papaja: Create APA manuscripts with R Markdown*. (2020).

Supplementary Material

Supplementary Methods

Calculating global geographic and linguistic proximity matrices.

Geographic distance between two nations was calculated as the logged geodesic distance between country capital cities (data from the R package *maps*; Brownrigg, 2018) using the R package *geosphere* (Hijmans, 2019). The geographic proximity matrix was computed as one minus the log geographic distance matrix scaled between 0 and 1.

Linguistic proximity between two nations was calculated as the cultural proximity between all languages spoken within those nations, weighted by speaker percentages. We acquired cultural proximity data by combining the language family trees provided by Glottolog v3.0 (Hammarström et al., 2017) into one global language tree (undated and unresolved). We calculated cultural proximity s between two languages j and k as the distance (in number of nodes traversed) of their most recent common ancestor i to the root of the tree, through the formula:

$$s_{jk} = \frac{n_r - n_i}{n_r}$$

where n_r is the maximum path length (in number of nodes traversed) leading to the pan-human root r, and n_i is the maximum path length leading to node i. We then combined these proximities with speaker data from Ethnologue 21 (Ethnologue, 2018) and compared every language spoken within those nations by at least 1 permille of the population, weighted by speaker percentages, through the formula:

$$w_{lm} = \Sigma\Sigma p_{lj} p_{mk} s_{jk}$$

where p_{lj} is the percentage of the population in nation l speaking language j, p_{mk} is the percentage of the population in nation m speaking language k, and s_{jk} is the proximity measure between languages j and k (Eff, 2008). The resulting linguistic proximity matrix was also scaled between 0 and 1 before analysis.

Bayesian models for reanalysis. We provide model formulae for our reanalyses of cross-national correlations, for a general bivariate case with standardised outcome Y and predictor X variables. In the naive regression model without controls for non-independence:

$$
\begin{aligned}
Y_i &\sim \text{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \beta X_i \\
\alpha &\sim \text{Normal}(0, 0.4) \\
\beta &\sim \text{Normal}(0, 0.4) \\
\sigma &\sim \text{Exponential}(5)
\end{aligned}
$$

The priors in this model were arrived at by prior predictive checks, with wider priors making predictions beyond the scale of standardised outcome variables and narrower priors being too informative.

To control for spatial non-independence, we add a Gaussian process to this model and feed it a scaled geographic distance matrix D based on Euclidean distances between latitude and longitude coordinates. This distance matrix is computed internally by the R package *brms* (Bürkner, 2017). The Gaussian process uses an exponentiated quadratic covariance kernel, the only covariance kernel currently supported by *brms*. The model formula is:

$$
\begin{aligned}
Y_i &\sim \text{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \kappa_{\text{NATION}[i]} + \beta X_i \\
\begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_3 \\ \ldots \\ \kappa_n \end{pmatrix} &\sim \text{MVNormal}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \ldots \\ 0 \end{pmatrix}, \mathbf{K} \right) \\
\mathbf{K}_{ij} &= sdgp^2 \exp\Big(-D_{ij}^2 / (2 lscale^2) \Big) \\
\alpha &\sim \text{Normal}(0, 0.4) \\
\beta &\sim \text{Normal}(0, 0.4) \\
\sigma &\sim \text{Exponential}(5) \\
sdgp &\sim \text{Exponential}(5) \\
lscale &\sim \text{InverseGamma}(?, ?)
\end{aligned}
$$

where n is the number of nations, and D_{ij}^2 reflects the squared Euclidean distances between latitude and longitude coordinates for the i-th and j-th nations. Notice that the inverse gamma prior on $lscale$ is left undetermined. This is because the *brms* package intelligently tunes the prior for this parameter based on the covariates of the Gaussian process (see https://betanalpha.github.io/assets/case_studies/gp_part3/part3.html).

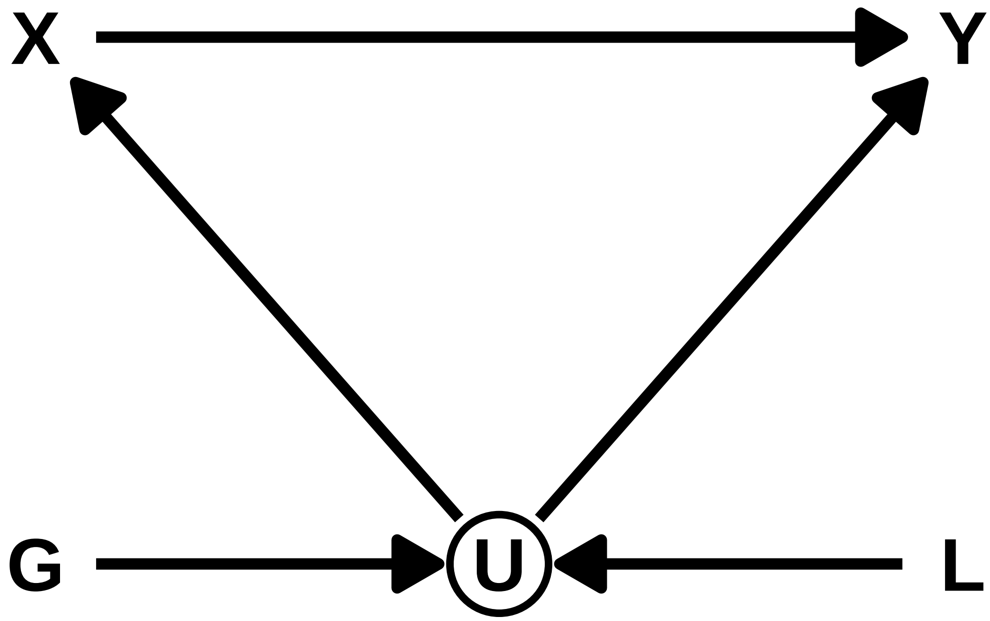
To control for cultural phylogenetic non-independence, we manually specify the covariance structure for nation random intercepts using a pre-computed linguistic proximity matrix L (see previous section). The covariance between two nations is assumed to be linearly proportional to the linguistic proximity between those nations. This assumption is justified if we assume that cultural traits evolve via Brownian motion along a language phylogeny. The non-centered parameterisation of this model is:

$$
\begin{aligned}
Y_i &\sim \text{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + z_{\text{NATION}[i]}\sigma_\alpha L + \beta X_i \\
\alpha &\sim \text{Normal}(0, 0.4) \\
\beta &\sim \text{Normal}(0, 0.4) \\
z_j &\sim \text{Normal}(0, 1) \\
\sigma_\alpha &\sim \text{Exponential}(5) \\
\sigma &\sim \text{Exponential}(5)
\end{aligned}
$$

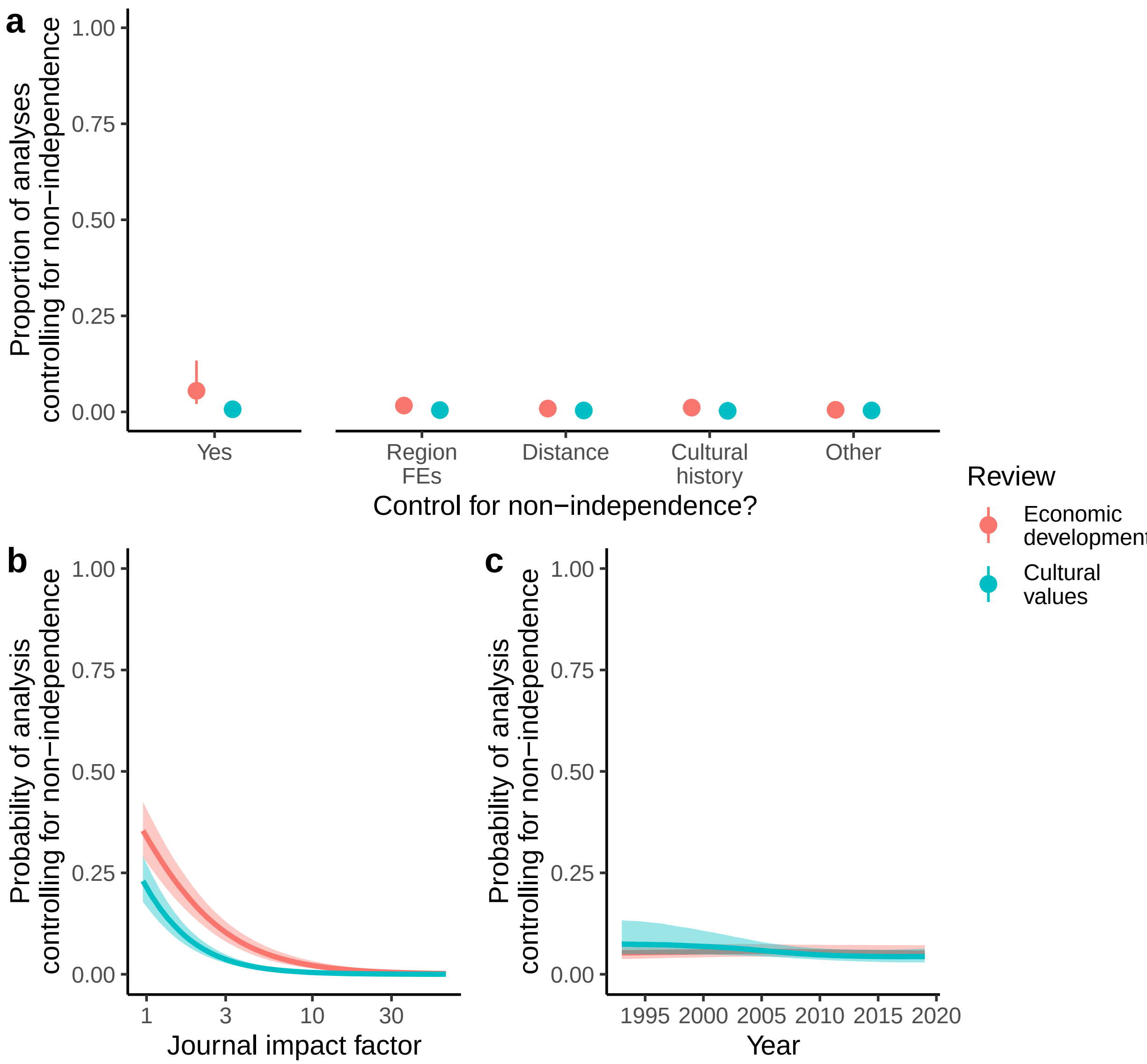
Finally, we can control for spatial and cultural phylogenetic non-independence simultaneously by including both a Gaussian process over latitude and longitude coordinates *and* nation random intercepts that covary according to linguistic proximity. The resulting model is as follows:

$$
\begin{aligned}
Y_i &\sim \text{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \kappa_{\text{NATION}[i]} + z_{\text{NATION}[i]}\sigma_\alpha L + \beta X_i \\
\begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_3 \\ \dots \\ \kappa_n \end{pmatrix} &\sim \text{MVNormal}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}, \mathbf{K} \right) \\
\mathbf{K}_{ij} &= sdgp^2 \exp\Big(- D_{ij}^2 / (2 lscale^2) \Big) \\
\alpha &\sim \text{Normal}(0, 0.4) \\
\beta &\sim \text{Normal}(0, 0.4) \\
z_j &\sim \text{Normal}(0, 1) \\
\sigma_\alpha &\sim \text{Exponential}(5) \\
\sigma &\sim \text{Exponential}(5) \\
sdgp &\sim \text{Exponential}(5) \\
lscale &\sim \text{InverseGamma}(?, ?)
\end{aligned}
$$

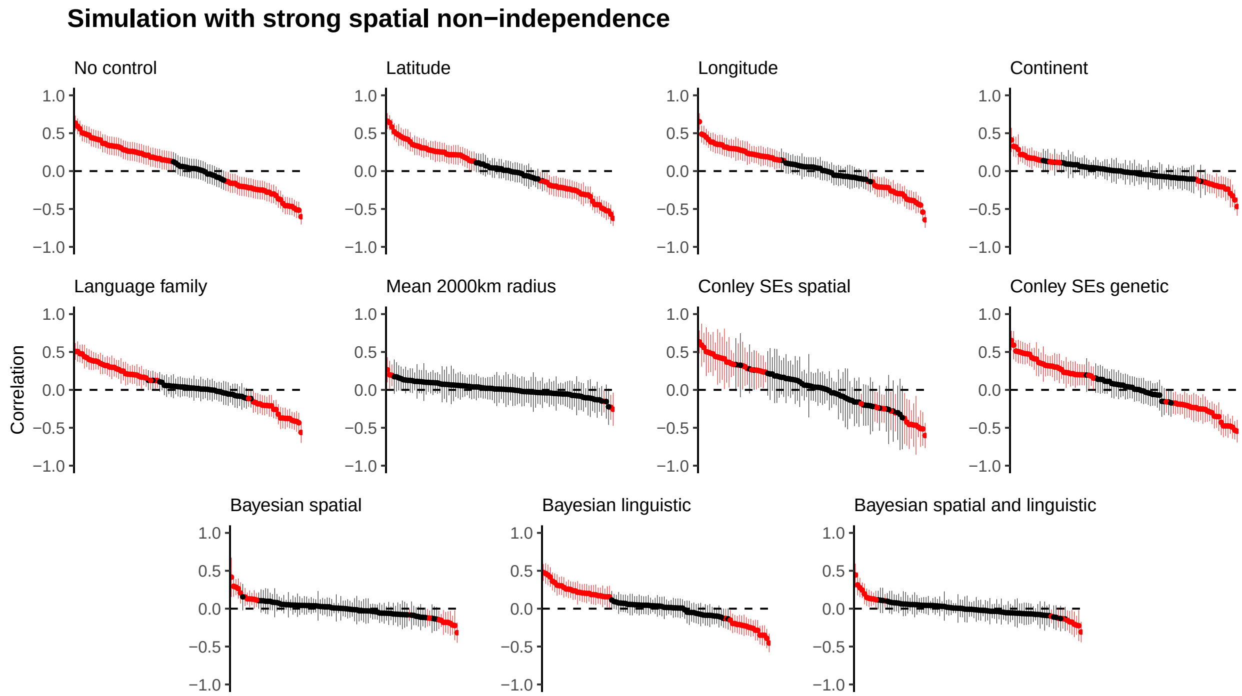
Supplementary Figures



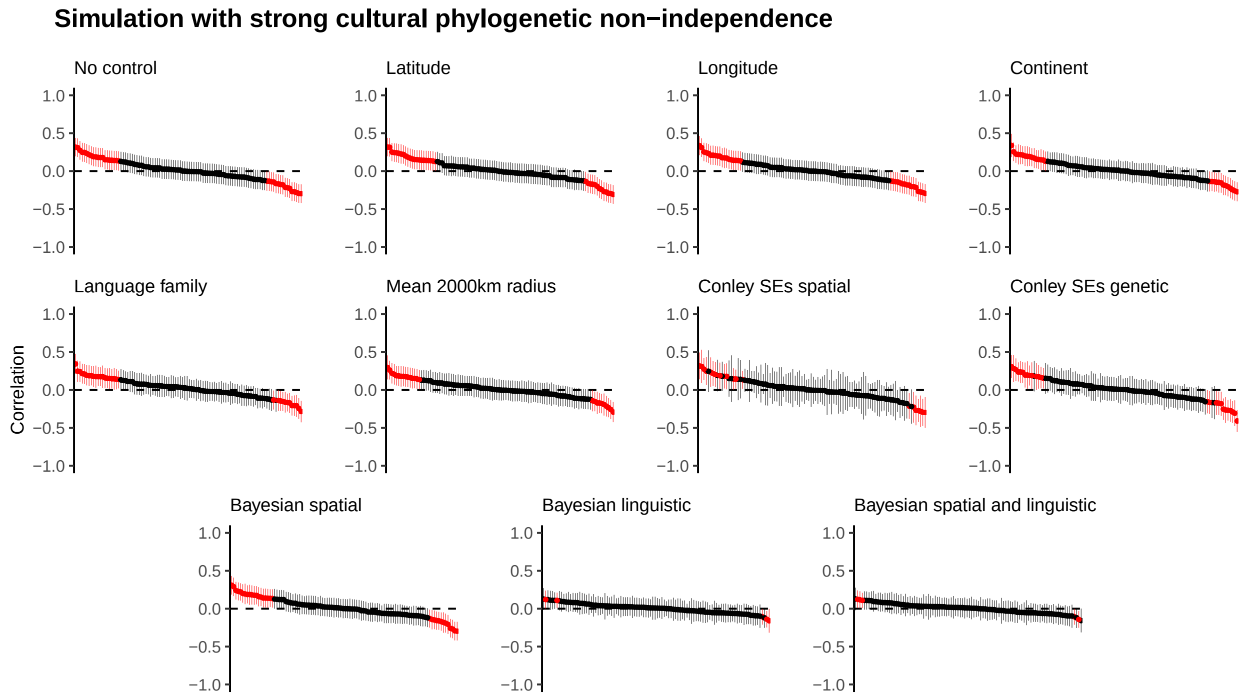
Supplementary Figure S1. A causal directed acyclic graph of spatial and cultural phylogenetic non-independence in cross-national studies. We are interested in estimating the direct effect of national-level exposure X on national-level outcome Y. But these variables are confounded by their common unobserved cause U. U is a stand-in for shared environmental, ecological, and geographic causes (e.g. climate, biodiversity, physical topography) and cultural and institutional causes (e.g. cultural norms, political systems). In this causal model, we need to condition on U to estimate the direct path from X to Y, but we cannot since it is unobserved. However, geographic G and linguistic L relationships between societies influence U, since changing a nation's spatial distance to or shared cultural ancestry with other nations will change its environmental and cultural traits. We can thus use G and L to model the covariation between X and Y induced by U. Failing to do this and simply estimating the bivariate correlation between X and Y will produce spurious relationships and residuals that are spatially and culturally non-independent around the world.



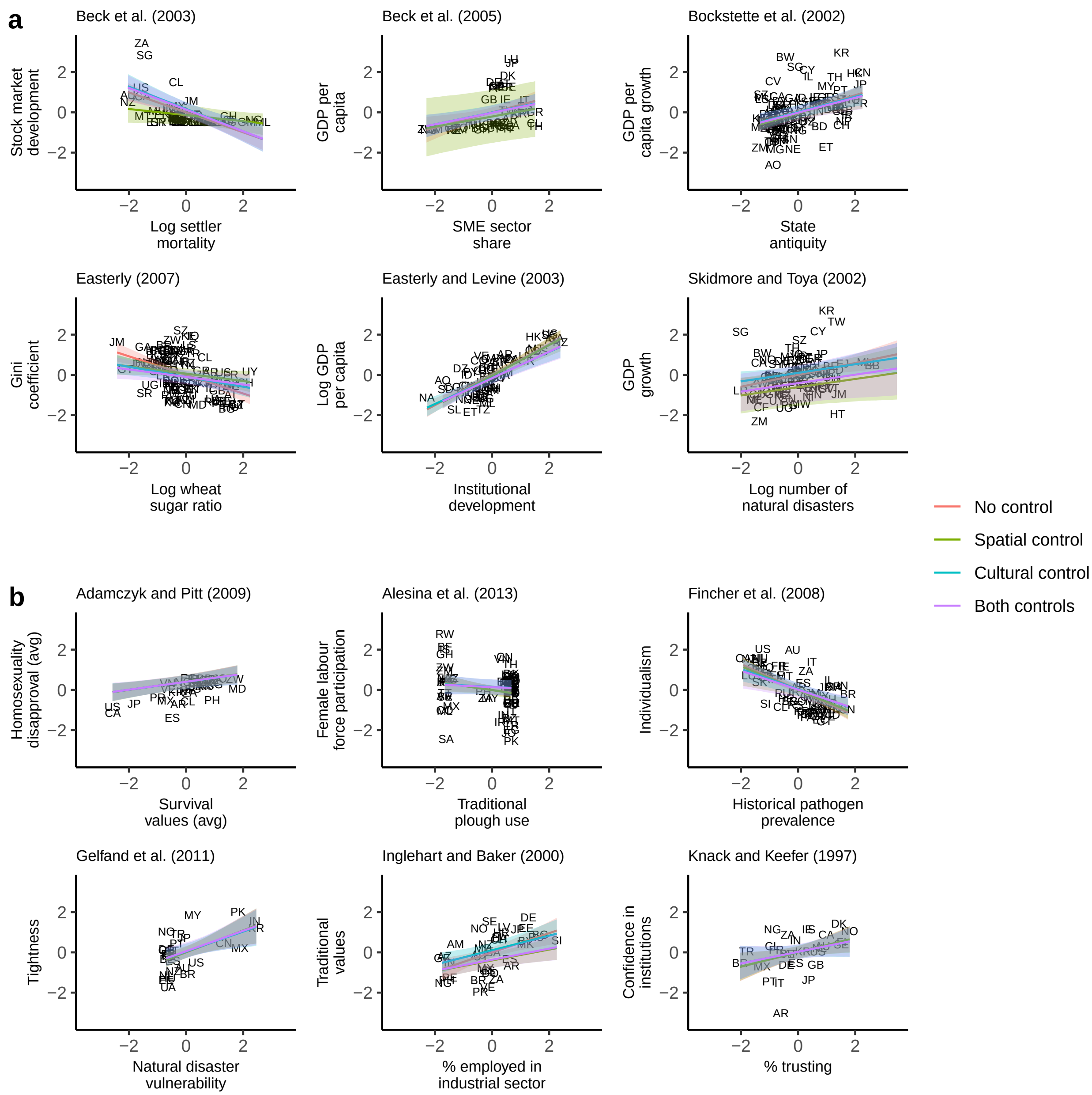
Supplementary Figure S2. Analysis-level results from systematic review of 100 highly-cited cross-national studies of economic development (red) and cultural values (blue). (a) Proportion of analyses accounting for non-independence, overall and split by common methods of controlling for non-independence. (b) The association between journal impact factor and the probability that an analysis accounts for non-independence. (c) Estimated trend over time for the probability that an analysis accounts for non-independence. Point ranges represent estimated proportions and 95% credible intervals. Lines and shaded areas are posterior median regression lines and 50% credible intervals from Bayesian multilevel models. Region FEs = region fixed effects.



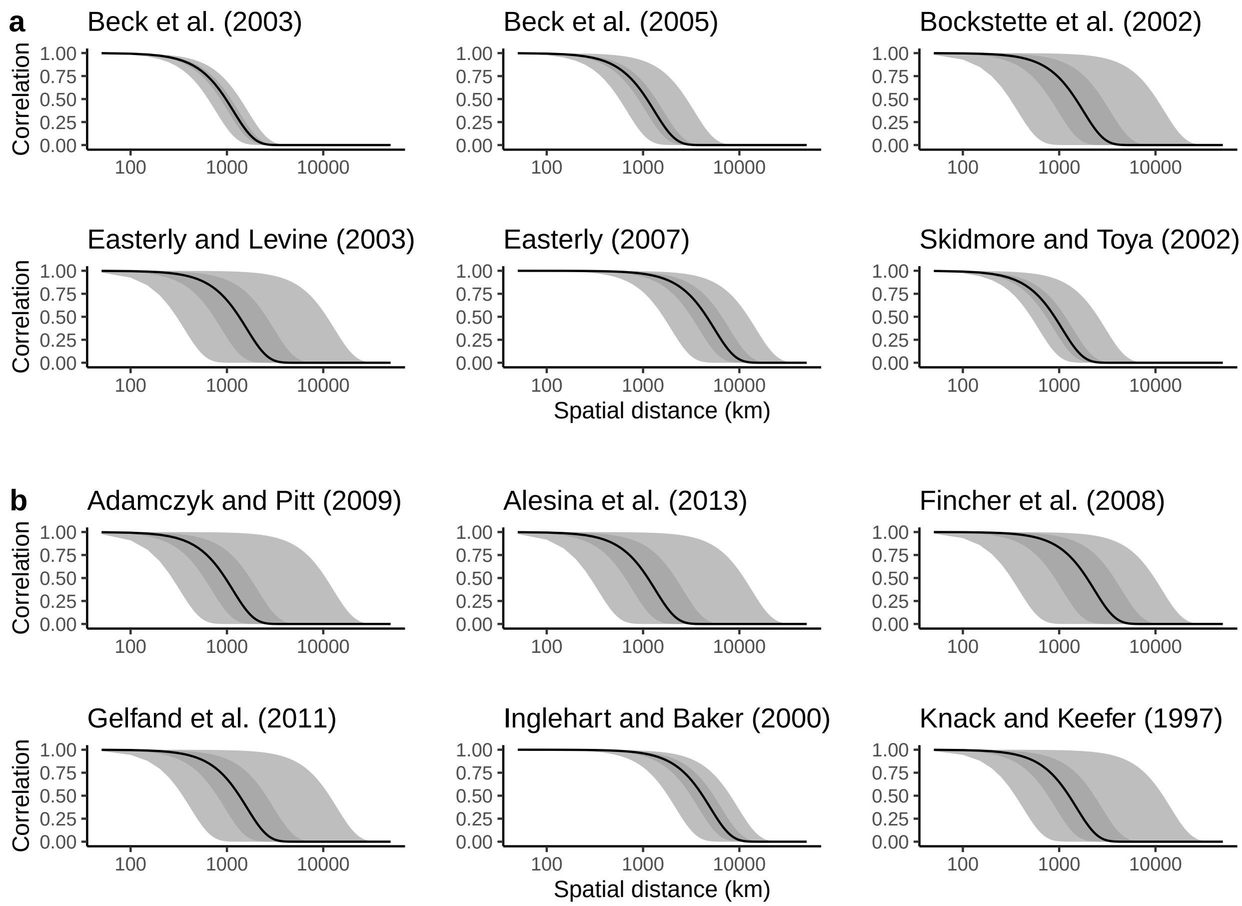
Supplementary Figure S3. Distribution of cross-national correlations from simulation study under strong spatial autocorrelation. In these simulations, the strength of spatial autocorrelation is set to 0.8 for both outcome and predictor variables. For frequentist regression models, point ranges represent correlation estimates and 95% confidence intervals. For Bayesian regression models, point ranges represent posterior means and 95% credible intervals. Correlations are ordered by effect size independently in each panel. Red point ranges indicate that the slope is “significant” (i.e. the 95% confidence / credible interval excludes zero). Black point ranges indicate that the slope is “not significant”.



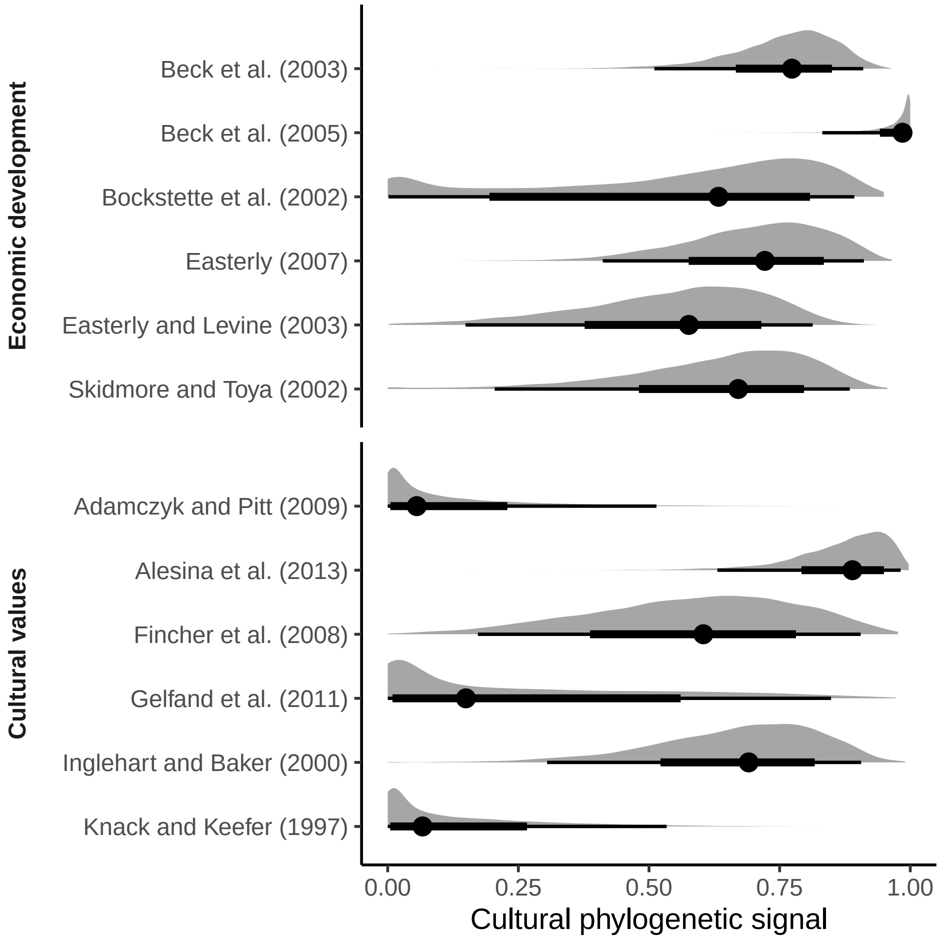
Supplementary Figure S4. *Distribution of cross-national correlations from simulation study under strong cultural phylogenetic autocorrelation.* In these simulations, the strength of cultural phylogenetic autocorrelation is set to 0.8 for both outcome and predictor variables. For frequentist regression models, point ranges represent correlation estimates and 95% confidence intervals. For Bayesian regression models, point ranges represent posterior means and 95% credible intervals. Correlations are ordered by effect size independently in each panel. Red point ranges indicate that the slope is "significant" (i.e. the 95% confidence / credible interval excludes zero). Black point ranges indicate that the slope is "not significant".



Supplementary Figure S5. Reanalysis models fitted to raw data, for economic development (a) and cultural values (b) studies. Data points are labelled using ISO 3166-1 alpha-2 letter country codes. In all reanalyses, outcome and predictor variables are standardised, making regression slopes comparable to Pearson's correlation coefficients. Lines and shaded areas represent posterior median regression lines and 95% credible intervals. For models with covariates (Adamzyck and Pitt 2009; Gelfand et al. 2011), marginal effects are presented holding all covariates at zero or their reference categories.



Supplementary Figure S6. Posterior estimates of Gaussian process functions mapping spatial autocorrelation onto geographic distance from our reanalyses of economic development (a) and cultural values (b) studies. Estimates are from models additionally controlling for cultural phylogenetic non-independence. The y-axis represents the amount of spatial autocorrelation between data points with increasing distance between those points on the x-axis (logged distance in kilometres). Lines and shaded areas represent median posterior spatial autocorrelation functions and 50% and 95% credible intervals.



Supplementary Figure S7. Posterior estimates of cultural phylogenetic signal from our re-analyses. Estimates are from models additionally controlling for spatial non-independence. Cultural phylogenetic signal is operationalised as the proportion of national-level variance explained by linguistic proximity between nations. Ridges are full posterior distributions, and points are posterior medians, and lines represent 50% and 95% credible intervals.

Supplementary Tables

Supplementary Table S1

Geographic and cultural phylogenetic signal results for economic development and cultural values variables. Signal estimates reflect the proportion of national-level variance explained by geographic and linguistic covariance matrices. Bayes Factors (BF) reflect support for the hypothesis that the signal estimate differs from zero. HDI = Human Development Index; GDPpc = gross domestic product per capita.

Outcome	Geographic signal	Cultural phylogenetic signal
HDI	0.37, 95% CI [0.18, 0.59], BF > 100	0.62, 95% CI [0.40, 0.80], BF > 100
GDPpc	0.42, 95% CI [0.20, 0.66], BF > 100	0.56, 95% CI [0.33, 0.78], BF > 100
GDPpc growth	0.65, 95% CI [0.09, 0.98], BF = 16.79	0.26, 95% CI [0.00, 0.70], BF = 1.16
Gini index	0.74, 95% CI [0.48, 0.97], BF > 100	0.25, 95% CI [0.02, 0.51], BF = 3.97
Traditional values	0.44, 95% CI [0.17, 0.76], BF > 100	0.54, 95% CI [0.23, 0.79], BF > 100
Survival values	0.20, 95% CI [0.01, 0.45], BF = 1.96	0.78, 95% CI [0.53, 0.96], BF > 100
Tightness	0.09, 95% CI [0.00, 0.32], BF = 0.33	0.84, 95% CI [0.61, 0.97], BF > 100
Individualism	0.26, 95% CI [0.00, 0.66], BF = 1.97	0.69, 95% CI [0.29, 0.96], BF > 100

Supplementary Table S2

List of 100 papers included in literature review, sorted by annual rate of citations since publication.

Review	Reference	Citations per year
Cultural values	Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. The Quarterly journal of economics, 112(4), 1251-1288.	121.50
	Inglehart, R., & Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. American sociological review, 19-51.	119.33
	Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., ... & Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. science, 332(6033), 1100-1104.	91.33
	Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. The quarterly journal of economics, 128(2), 469-530.	57.12
	Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. Journal of personality and social psychology, 94(1), 168.	49.15
	Hofstede, G., & McCrae, R. R. (2004). Personality and culture revisited: Linking traits and dimensions of culture. Cross-cultural research, 38(1), 52-88.	35.53
	Delhey, J., Newton, K., & Welzel, C. (2011). How general is trust in "most people"? Solving the radius of trust problem. American Sociological Review, 76(5), 786-807.	35.10

Table S2 continued

Review	Reference	Citations per year
	Fincher, C. L., Thornhill, R., Murray, D. R., & Schaller, M. (2008). Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. Proceedings of the Royal Society B: Biological Sciences, 275(1640), 1279-1285.	33.38
	Schneider, S. L. (2008). Anti-immigrant attitudes in Europe: Outgroup size and perceived ethnic threat. European Sociological Review, 24(1), 53-67.	30.54
	Franzen, A., & Meyer, R. (2010). Environmental attitudes in cross-national perspective: A multilevel analysis of the ISSP 1993 and 2000. European sociological review, 26(2), 219-234.	30.27
	Steenkamp, J. B. E., Ter Hofstede, F., & Wedel, M. (1999). A cross-national investigation into the individual and national cultural antecedents of consumer innovativeness. Journal of marketing, 63(2), 55-69.	29.32
	Paxton, P. (2002). Social capital and democracy: An interdependent relationship. American sociological review, 254-277.	29.21
	Santos, H. C., Varnum, M. E., & Grossmann, I. (2017). Global increases in individualism. Psychological science, 28(9), 1228-1239.	28.25
	Van Zanten, B. T., Van Berkel, D. B., Meentemeyer, R. K., Smith, J. W., Tieskens, K. F., & Verburg, P. H. (2016). Continental-scale quantification of landscape values using social media data. Proceedings of the National Academy of Sciences, 113(46), 12974-12979.	27.00
	Beugelsdijk, S., & Welzel, C. (2018). Dimensions and dynamics of national culture: Synthesizing Hofstede with Inglehart. Journal of cross-cultural psychology, 49(10), 1469-1505.	24.00

Table S2 continued

Review	Reference	Citations per year
	Stephan, U., & Uhlaner, L. M. (2010). Performance-based vs socially supportive culture: A cross-national study of descriptive norms and entrepreneurship. Journal of International Business Studies, 41(8), 1347-1364.	23.91
	Dakhli, M., & De Clercq, D. (2004). Human capital, social capital, and innovation: a multi-country study. Entrepreneurship & regional development, 16(2), 107-128.	23.53
	Adamczyk, A., & Pitt, C. (2009). Shaping attitudes about homosexuality: The role of religion and cultural context. Social Science Research, 38(2), 338-351.	21.75
	Mayda, A. M., & Rodrik, D. (2005). Why are some people (and countries) more protectionist than others?. European Economic Review, 49(6), 1393-1430.	21.69
	Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. Journal of Cross-cultural psychology, 36(2), 264-277.	20.62
	Tam, K. P., & Chan, H. W. (2018). Generalized trust narrows the gap between environmental concern and pro-environmental behavior: Multilevel evidence. Global Environmental Change, 48, 182-194.	20.00
	Kuppens, P., Realo, A., & Diener, E. (2008). The role of positive and negative emotions in life satisfaction judgment across nations. Journal of personality and social psychology, 95(1), 66.	19.62

Table S2 continued

Review	Reference	Citations per year
	Tam, K. P., & Chan, H. W. (2017). Environmental concern has a weaker association with pro-environmental behavior in some societies than others: A cross-cultural psychology perspective. Journal of Environmental Psychology, 53, 213-223.	18.75
	Miska, C., Szőcs, I., & Schiffinger, M. (2018). Culture's effects on corporate sustainability practices: A multi-domain and multi-level view. Journal of World Business, 53(2), 263-279.	17.67
	Oishi, S., Diener, E., Lucas, R. E., & Suh, E. M. (2009). Cross-cultural variations in predictors of life satisfaction: Perspectives from needs and values. In Culture and well-being (pp. 109-127). Springer, Dordrecht.	16.00
	Malka, A., Soto, C. J., Inzlicht, M., & Lelkes, Y. (2014). Do needs for security and certainty predict cultural and economic conservatism? A cross-national analysis. Journal of personality and social psychology, 106(6), 1031.	16.00
	Eom, K., Kim, H. S., Sherman, D. K., & Ishii, K. (2016). Cultural variability in the link between environmental concern and support for environmental action. Psychological Science, 27(10), 1331-1339.	15.80
	Larsen, C. A. (2008). The institutional logic of welfare attitudes: How welfare regimes influence public support. Comparative political studies, 41(2), 145-168.	15.62
	Sortheix, F. M., & Schwartz, S. H. (2017). Values that underlie and undermine well–being: Variability across countries. European Journal of Personality, 31(2), 187-201.	15.50

Table S2 continued

Review	Reference	Citations per year
	Van der Meer, T., & Hakhverdian, A. (2017). Political trust as the evaluation of process and performance: A cross-national study of 42 European countries. Political Studies, 65(1), 81-102.	15.25
	Resick, C. J., Hanges, P. J., Dickson, M. W., & Mitchelson, J. K. (2006). A cross-cultural examination of the endorsement of ethical leadership. Journal of Business Ethics, 63(4), 345-359.	15.20
	Murray, D. R., & Schaller, M. (2010). Historical prevalence of infectious diseases within 230 geopolitical regions: A tool for investigating origins of culture. Journal of Cross-Cultural Psychology, 41(1), 99-108.	15.00
	Welzel, C., Inglehart, R., & Kligemann, H. D. (2003). The theory of human development: A cross-cultural analysis. European Journal of Political Research, 42(3), 341-379.	14.94
	Löckenhoff, C. E., De Fruyt, F., Terracciano, A., McCrae, R. R., De Bolle, M., Costa, P. T., ... & Yik, M. (2009). Perceptions of aging across 26 cultures and their culture-level associates. Psychology and aging, 24(4), 941.	14.75
	Fulmer, C. A., Gelfand, M. J., Kruglanski, A. W., Kim-Prieto, C., Diener, E., Pierro, A., & Higgins, E. T. (2010). On "feeling right" in cultural contexts: How person-culture match affects self-esteem and subjective well-being. Psychological Science, 21(11), 1563-1569.	14.64
	Gifford, R., Scannell, L., Kormos, C., Smolova, L., Biel, A., Boncu, S., ... & Uzzell, D. (2009). Temporal pessimism and spatial optimism in environmental assessments: An 18-nation study. Journal of environmental psychology, 29(1), 1-12.	14.25

Table S2 continued

Review	Reference	Citations per year
	Tranter, B., & Booth, K. (2015). Scepticism in a changing climate: A cross-national study. Global Environmental Change, 33, 154-164.	14.17
	Pisano, I., & Lubell, M. (2017). Environmental behavior in cross-national perspective: A multilevel analysis of 30 countries. Environment and Behavior, 49(1), 31-58.	13.75
	Kesler, C., & Bloemraad, I. (2010). Does immigration erode social capital? The conditional effects of immigration-generated diversity on trust, membership, and participation across 19 countries, 1981–2000. Canadian Journal of Political Science/Revue canadienne de science politique, 43(2), 319-347.	13.73
	Gelissen, J. (2007). Explaining popular support for environmental protection: A multilevel analysis of 50 nations. Environment and behavior, 39(3), 392-415.	13.71
	Thorisdottir, H., Jost, J. T., Liviatan, I., & Shrout, P. E. (2007). Psychological needs and values underlying left-right political orientation: Cross-national evidence from Eastern and Western Europe. Public Opinion Quarterly, 71(2), 175-203.	13.50
	Hörisch, J., Kollat, J., & Brieger, S. A. (2017). What influences environmental entrepreneurship? A multilevel analysis of the determinants of entrepreneurs' environmental orientation. Small Business Economics, 48(1), 47-69.	13.50
	Fungáčová, Z., Hasan, I., & Weill, L. (2019). Trust in banks. Journal of Economic Behavior & Organization, 157, 452-476.	13.50

Table S2 continued

Review	Reference	Citations per year
	Cuddy, A. J., Wolf, E. B., Glick, P., Crotty, S., Chong, J., & Norton, M. I. (2015). Men as cultural ideals: Cultural values moderate gender stereotype content. Journal of personality and social psychology, 109(4), 622.	13.17
	Mata, R., Josef, A. K., & Hertwig, R. (2016). Propensity for risk taking across the life span and around the globe. Psychological science, 27(2), 231-243.	13.00
	Venaik, S., & Brewer, P. (2010). Avoiding uncertainty in Hofstede and GLOBE. Journal of international business studies, 41(8), 1294-1315.	12.55
	Schimmack, U., Oishi, S., & Diener, E. (2005). Individualism: A valid and important dimension of cultural differences between nations. Personality and Social Psychology Review, 9(1), 17-31.	12.31
	Kvaløy, B., Finseraas, H., & Listhaug, O. (2012). The publics' concern for global warming: A cross-national study of 47 countries. Journal of Peace Research, 49(1), 11-22.	12.00
	Mikucka, M., Sarracino, F., & Dubrow, J. K. (2017). When does economic growth improve life satisfaction? Multilevel analysis of the roles of social trust and income inequality in 46 countries, 1981–2012. World Development, 93, 447-459.	12.00
	Miyamoto, Y., Yoo, J., Levine, C. S., Park, J., Boylan, J. M., Sims, T., ... & Ryff, C. D. (2018). Culture and social hierarchy: Self- and other-oriented correlates of socioeconomic status across cultures. Journal of Personality and Social Psychology, 115(3), 427.	12.00

Table S2 continued

Review	Reference	Citations per year
Economic development	Inglehart, R., & Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. American sociological review, 19-51.	119.33
	King, R. G., & Levine, R. (1993). Finance and growth: Schumpeter might be right. The quarterly journal of economics, 108(3), 717-737.	90.00
	Treisman, D. (2000). The causes of corruption: a cross-national study. Journal of public economics, 76(3), 399-457.	73.67
	Selden, T. M., & Song, D. (1994). Environmental quality and development: is there a Kuznets curve for air pollution emissions?. Journal of Environmental Economics and management, 27(2), 147-162.	49.07
	Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. Journal of Monetary economics, 34(2), 143-173.	47.89
	Acemoglu, D., Johnson, S., Robinson, J. A., & Yared, P. (2008). Income and democracy. American Economic Review, 98(3), 808-42.	44.62
	Easterly, W., & Levine, R. (2003). Tropics, germs, and crops: how endowments influence economic development. Journal of monetary economics, 50(1), 3-39.	36.10
	Wimmer, A., Cederman, L. E., & Min, B. (2009). Ethnic politics and armed conflict: A configurational analysis of a new global data set. American sociological review, 74(2), 316-337.	34.42

Table S2 continued

Review	Reference	Citations per year
	Redding, S., & Venables, A. J. (2004). Economic geography and international inequality. Journal of international Economics, 62(1), 53-82.	32.65
	Liobikienė, G., Mandravickaitė, J., & Bernatonienė, J. (2016). Theory of planned behavior approach to understand the green purchasing behavior in the EU: A cross-cultural study. Ecological Economics, 125, 38-46.	31.00
	McCrae, R. R., & Terracciano, A. (2005). Personality profiles of cultures: aggregate personality traits. Journal of personality and social psychology, 89(3), 407.	26.12
	Faccio, M. (2010). Differences between politically connected and nonconnected firms: A cross-country analysis. Financial management, 39(3), 905-928.	25.73
	Ginarte, J. C., & Park, W. G. (1997). Determinants of patent rights: A cross-national study. Research policy, 26(3), 283-301.	25.08
	Beugelsdijk, S., & Welzel, C. (2018). Dimensions and dynamics of national culture: Synthesizing Hofstede with Inglehart. Journal of cross-cultural psychology, 49(10), 1469-1505.	24.00
	Jorgenson, A. K., & Clark, B. (2012). Are the economy and the environment decoupling? A comparative international study, 1960–2005. American Journal of Sociology, 118(1), 1-44.	23.00
	Durlauf, S. N., & Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour. Journal of applied econometrics, 10(4), 365-384.	22.88

Table S2 continued

Review	Reference	Citations per year
	Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. Journal of development economics, 84(2), 755-776.	21.93
	Fuwa, M. (2004). Macro-level gender inequality and the division of household labor in 22 countries. American sociological review, 69(6), 751-767.	21.24
	Smith, M. D., Rabbitt, M. P., & Coleman-Jensen, A. (2017). Who are the world's food insecure? New evidence from the Food and Agriculture Organization's food insecurity experience scale. World Development, 93, 402-412.	20.75
	Beck, T., Demirgüç-Kunt, A., & Levine, R. (2003). Law, endowments, and finance. Journal of financial Economics, 70(2), 137-181.	20.28
	Schreinemachers, P., & Tipraqsa, P. (2012). Agricultural pesticides and land use intensification in high, middle and low income countries. Food policy, 37(6), 616-626.	20.11
	Poe, S. C., Tate, C. N., & Keith, L. C. (1999). Repression of the human right to personal integrity revisited: A global cross-national study covering the years 1976–1993. International studies quarterly, 43(2), 291-313.	19.36
	Bloom, D. E., Sachs, J. D., Collier, P., & Udry, C. (1998). Geography, demography, and economic growth in Africa. Brookings papers on economic activity, 1998(2), 207-295.	18.61
	Skidmore, M., & Toya, H. (2002). Do natural disasters promote long-run growth?. Economic inquiry, 40(4), 664-687.	18.47

Table S2 continued

Review	Reference	Citations per year
	Kurtz, M. J., & Schrank, A. (2007). Growth and governance: Models, measures, and mechanisms. The Journal of Politics, 69(2), 538-554.	17.53
	Klasen, S., & Lamanna, F. (2009). The impact of gender inequality in education and employment on economic growth: new evidence for a panel of countries. Feminist economics, 15(3), 91-132.	17.25
	Welzel, C., Inglehart, R., & Kligemann, H. D. (2003). The theory of human development: A cross-cultural analysis. European Journal of Political Research, 42(3), 341-379.	14.94
	Brülhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. Journal of Urban Economics, 65(1), 48-63.	14.33
	Beck, T., Demirguc-Kunt, A., & Levine, R. (2005). SMEs, growth, and poverty: Cross-country evidence. Journal of economic growth, 10(3), 199-229.	14.12
	Ehrhardt-Martinez, K., Crenshaw, E. M., & Jenkins, J. C. (2002). Deforestation and the environmental Kuznets curve: A cross-national investigation of intervening mechanisms. Social Science Quarterly, 83(1), 226-243.	13.05
	Rai, D., Zitko, P., Jones, K., Lynch, J., & Araya, R. (2013). Country-and individual-level socioeconomic determinants of depression: multilevel cross-national comparison. The British Journal of Psychiatry, 202(3), 195-203.	12.75
	Holzner, M. (2011). Tourism and economic development: The beach disease?. Tourism Management, 32(4), 922-933.	12.70

Table S2 continued

Review	Reference	Citations per year
	Calderón, C., Moral-Benito, E., & Servén, L. (2015). Is infrastructure capital productive? A dynamic heterogeneous approach. Journal of Applied Econometrics, 30(2), 177-198.	12.67
	Roskam, A. J. R., Kunst, A. E., Van Oyen, H., Demarest, S., Klumbiene, J., Regidor, E., ... & additional participants to the study. (2010). Comparative appraisal of educational inequalities in overweight and obesity among adults in 19 European countries. International journal of epidemiology, 39(2), 392-404.	12.64
	Curtis, J. E., Baer, D. E., & Grabb, E. G. (2001). Nations of joiners: Explaining voluntary association membership in democratic societies. American Sociological Review, 783-805.	12.40
	Adams, S., & Klobodu, E. K. M. (2018). Financial development and environmental degradation: does political regime matter?. Journal of Cleaner Production, 197, 1472-1479.	12.33
	Qian, Y. (2007). Do national patent laws stimulate domestic innovation in a global patenting environment? A cross-country analysis of pharmaceutical patent protection, 1978–2002. The Review of Economics and Statistics, 89(3), 436-453.	12.29
	Gaddis, I., & Klasen, S. (2014). Economic development, structural change, and women's labor force participation. Journal of Population Economics, 27(3), 639-681.	11.86
	Bockstette, V., Chanda, A., & Putterman, L. (2002). States and markets: The advantage of an early start. Journal of Economic growth, 7(4), 347-369.	11.84

Table S2 continued

Review	Reference	Citations per year
	Levels, M., Dronkers, J., & Kraaykamp, G. (2008). Immigrant children's educational achievement in western countries: origin, destination, and community effects on mathematical performance. American Sociological Review, 73(5), 835-853.	11.62
	Klasen, S. (2002). Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development. The World Bank Economic Review, 16(3), 345-373.	11.42
	Knudsen, K., & Wærness, K. (2008). National context and spouses' housework in 34 countries. European Sociological Review, 24(1), 97-113.	11.15
	Rigobon, R., & Rodrik, D. (2005). Rule of law, democracy, openness, and income: Estimating the interrelationships. Economics of transition, 13(3), 533-564.	11.00
	Beladi, H., Chao, C. C., Ee, M. S., & Hollas, D. (2019). Does medical tourism promote economic growth? A cross-country analysis. Journal of Travel Research, 58(1), 121-135.	11.00
	Roe, M. J., & Siegel, J. I. (2011). Political instability: Effects on financial development, roots in the severity of economic inequality. Journal of Comparative Economics, 39(3), 279-309.	10.90
	Roberts, J. T., & Grimes, P. E. (1997). Carbon intensity and economic development 1962–1991: A brief exploration of the environmental Kuznets curve. World development, 25(2), 191-198.	10.71

Table S2 continued

Review	Reference	Citations per year
	Reinsberg, B., Kentikelenis, A., Stubbs, T., & King, L. (2019). The world system and the hollowing out of state capacity: How structural adjustment programs affect bureaucratic quality in developing countries. American Journal of Sociology, 124(4), 1222-1257.	10.50
	Jones-Smith, J. C., Gordon-Larsen, P., Siddiqi, A., & Popkin, B. M. (2011). Cross-national comparisons of time trends in overweight inequality by socioeconomic status among women using repeated cross-sectional surveys from 37 developing countries, 1989–2007. American journal of epidemiology, 173(6), 667-675.	10.40
	Demirgüç-Kunt, A., & Levine, R. (1996). Stock market development and financial intermediaries: stylized facts. The World Bank Economic Review, 10(2), 291-321.	10.36
	Sethi, D., & Acharya, D. (2018). Financial inclusion and economic growth linkage: Some cross country evidence. Journal of Financial Economic Policy.	10.33

Supplementary References

Brownrigg, R. *maps: Draw geographical maps.* (2018).

Bürkner, P.-C. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* **80**, 1—28 (2017).

Eff, E. A. Weight matrices for cultural proximity: Deriving weights from a language phylogeny. *Structure and Dynamics* **3**, (2008).

Ethnologue: Languages of the world. (SIL International, 2018).

Hammarström, H., Forkel, R., Haspelmath, M. & Bank, S. *Glottolog 3.0.* (Max Planck Institute for the Science of Human History, 2017). doi: 10.5281/zenodo.4061162

Hijmans, R. J. *geosphere: Spherical trigonometry.* (2019).