The Sustainable Development Oxymoron: Quantifying and Modelling the Incompatibility of Sustainable Development Goals

Viktoria Spaisera, Shyam Ranganathanb, Ranjula Bali Swanic,d, David J.T. Sumpter
e
aSchool of Politics and International Studies, University of Leeds, Leeds, LS2 9JT, UK
bDepartment of Statistics, Virginia Tech, Blacksburg, VA 24061, USA
cSchool of Social Sciences, Södertörn University, Stockholm, 141 89, Sweden
dMISUM, Stockholm School of Economics, Stockholm, 113 83, Sweden
eDepartment of Mathematics, Uppsala University, Uppsala, 751 06, Sweden

Abstract

In 2015, the UN adopted a new set of Sustainable Development Goals (SDGs) to eradicate poverty, establish socio-economic inclusion and protect the environment. Critical voices such as the International Council for Science, however, have expressed concerns about potential incompatibility of the SDGs, specifically the incompatibility of socio-economic development and environmental sustainability. In this paper we test, quantify and model the alleged inconsistency of SDGs. Our analyses show which SDGs are consistent and which conflicting. We measure the extent of inconsistency and conclude that the SDG agenda will fail as a whole if we continue with business as usual. We further explore the nature of the inconsistencies using dynamical systems models, which reveal that the focus on economic growth and consumption as a means for development underlies the inconsistency. Our models also show that there are factors which can contribute to development (health programs, government investment) on the one hand and ecological sustainability (renewable energy) on the other, without triggering the conflict between incompatible SDGs.

Keywords: Sustainable Development Goals, sustainable development incompatibility, dynamical systems modeling, factor analysis, feature selection, UN data revolution

1. Introduction

The concept of sustainable development was conceived in 1987 by the Brundtland Commission as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (World Commission on Environment and Development, 1987). Since then it has been repeatedly challenged with the argument that economic growth is not sustainable because it is accompanied by a depletion of natural resources and deterioration of environmental service (Repetto et al., 1989; Pearce and Atkinson, 1993; Hamilton and Clemens, 1999). Redclift (2005) even called sustainable development an oxymoron that disguises the inherent conflict of human and natural systems. And Dasgupta (2013) noted that “the entire architecture of contemporary development thinking is stacked against nature” (p.2). Despite this challenge never being addressed, the world now prepares to take up the ambitious Sustainable Development Goals (SDG). The SDGs go beyond the earlier Millennium Development Goals (MDG) program in that, as well as addressing extreme poverty, they also focus on socio-economic inclusion and ecological sustainability. Seventeen SDG goals (Figure 1) have been defined by the UN for the next 15 years, with both developing and developed countries held accountable for meeting the goals (Leadership Council of the Sustainable Development Solutions Network, 2015; United National Department of Economic and Social Affairs, 2015).

The International Council for Science (ICSU) reviewed the SDGs and criticized “that the framework as a whole might not be internally consistent – and as a result not be sustainable”, if interlinkages, complex dynamics and conflicting relations between the goals are ignored (International Council for Science and International Social Science Council, 2015, p.9). The conflict between ecological sustainability and socio-economic progression is of particular concern to the ICSU. However, the ICSU does not provide empirical evidence for its concerns. In fact, little is known about the nature and extent of the repeatedly claimed incompatibility of sustainability and development (Stern et al., 1996; Redclift, 2002, 2005; Luke, 2005; Brown, 2015; Saboori and Sulaiman, 2013).
The aim of this study is to use the increasing amount of data available on development (Leadership Council of the Sustainable Development Solutions Network, 2015) to investigate potential contradictions in the SDGs. We compiled an extensive, cross-country time series dataset, including data from the World Bank, Polity IV, CIRI Human Rights Data Project, Freedom House and the Heritage Foundation/The Wall Street Journal. The compiled dataset consists of 1,423 economic, social, environmental and political indicators for 217 countries, covering the period 1980 to 2014 (see Appendix A1 Data). We first applied Confirmatory and Exploratory Factor Analyses (Harrington, 2008; Reinecke, 2005) to test and quantify the (in-)consistency of the SDGs. To then understand the dynamics and mechanisms of sustainable development inconsistencies, we use a Feature Selection Algorithm (Guyon and Elisseeff, 2003; Mehmood et al., 2012) to pre-select best predictors for sustainable development from the large set of potential predictors and then fit dynamical systems models (Ranganathan et al., 2014a) with the pre-selected predictors. Finally we suggest how the resulting dynamical systems models can be used to monitor sustainable development.

2. Data and Methods

To test, quantify and model the alleged inconsistency of the Sustainable Development Goals (SDGs), we used data from various sources. First we downloaded all the data provided by the World Bank for the period 1980-2014, through the World Bank Data API. This includes all World Development Indicators (WDI), African development Indicators (ADI), International Debt Statistics, Millennium Development Goals data, Education Statistics, Gender Statistics, Health, Nutrition and Population Statistics, Worldwide Governance Indicators (WGI), Climate Change Data, Global Financial Development data, Doing Business data, Enterprise Surveys data, Poverty and Equity Database, The Changing Wealth of Nations data, ILO (International Labour Organization) data and IDA (International Development Association) Results Measurement System (see Appendix A1 Data) (World Bank Development Data Group, 2014). To this World Bank dataset we added datasets from Polity IV containing measures on democracy vs. autocracy and regime durability (Marshall et al., 2014), CIRI Human Rights Data Project containing various human rights indices (Cingranelli et al., 2013), Centre for Systemic Peace data containing data on conflicts (interstate wars, civil wars, etc.), state fragility, governance and immigration (compiled by Monty G. Marshall, 2014), the Freedom House data containing the indices for political and civil rights as well as freedom of press (Freedom House, 2014a,b) and the Heritage Foundation/Wall Street Journal data containing measures on economic freedom, such as property rights, freedom from corruption, fiscal freedom, government spending, business freedom, labor freedom, trade freedom, investment freedom and financial freedom (Miller et al., 2014). Please see Appendix A1 Data for further details on the dataset.

We first applied Confirmatory and Exploratory Factor Analyses to test and quantify the inconsistency of the SDGs. A Confirmatory Factor Analysis (CFA) is a special case of a covariance-based structural equation model. The CFA assumes a theoretical or conceptual model with defined observed variables, which serve as indicators for the postulated latent factor. A linear combination of the observed variables and measurement errors gives the latent factor, which represents the theoretical construct. The CFA can be hierarchical involving first and second order latent constructs. The observed variables construct two or more first order latent factors and these first order latent factors then construct one or several overarching second order factors. Standardized factor loadings and z-scores, which describe the relation between the estimated value (factor loading) and its standard error, are used to evaluate the contribution of the observation variables to the latent factor. Model indices such as the Root Mean Square Error of Approximation (RMSEA), the Standardised Root Mean Square Residual (SRMR), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI) are used to assess the overall goodness of the CFA model (Harrington, 2008; Reinecke, 2005).

An Exploratory Factor Analysis (EFA) is used to test whether the correlation structure of the observed variables allows the extraction of one or several factors, and thus, to what extent the observed variables can be predicted by one or several latent factors (Thompson, 2004). Similar to a CFA, standardized factor loadings and z-scores are used to evaluate the contribution of the observation variables to the latent factor. The goodness of the EFA-extracted latent factor model is subsequently tested with a CFA approach.

Factor scores can be used to construct a latent variable. Typically, factors scores are linear combinations of the observed variables based on the shared variance between the observed variable and the factor (latent variable) and the variance that is not measured (Christine DiStefano and Mindrila, 2009). We used this approach to create the latent variable $L$, which consists of the three observable variables Child Mortality, Education (reverse coded, see Appendix A1.1), and CO2 emissions per capita.
To then understand the dynamics and mechanisms of sustainable development inconsistencies we used a Feature Selection Algorithm (Guyon and Elisseeff, 2003; Mehmood et al., 2012) to preselect best predictors for sustainable development from the large set of potential predictors. Specifically, we used the Variable Elimination Algorithm (Mehmood et al., 2012), a supervised feature selection machine learning method based on partial least squares regression. The uninformative variable elimination is a wrapper method, based on procedures iterating between model fitting and variable selection. The search algorithm extracts a subset of relevant variables and evaluates each subset by fitting a model. The variables can then be ranked based on an estimated relevance measure.

Finally, we fitted dynamical systems models (Ranganathan et al., 2014a) with the pre-selected predictors. We used the data to inform model selection from a large pool of potential models by fitting polynomial differential equations that are capable of capturing various linear and non-linear dynamical patterns in the data. For instance, changes in $L$ between times $t$ and $t+1$ were modeled as a function of all possible combinations of numerous pre-defined polynomial terms involving the pre-selected predictors at time $t$ with varying complexity (number of polynomial terms included in the model). Bayesian model selection as outlined in Ranganathan et al. (2014a) was used to identify the overall best-fitting model. Please see the Appendix A2 Methods for further details on our methodological approach.

3. Testing the Consistency of Sustainable Development

We tested the consistency of the Sustainable Development Goals (SDGs) using the Confirmatory Factor Analysis (CFA) approach (see Appendix A2 Methods). The UN SDG concept suggests a hierarchical latent structure. The first-order factors represent the three main pillars of the SDGs: to end poverty, ensure socio-economic inclusion, and protect the environment. The second-order, overarching factor represents sustainable development (Leadership Council of the Sustainable Development Solutions Network, 2015; United National Department of Economic and Social Affairs, 2015). Each of the three main pillars further break down in to 16 goals (plus the 17th administrative goal) which can be related to specific indicators identified by the UN (Leadership Council of the Sustainable Development Solutions Network, 2015) (see Figure 1 and Appendix A1 Data).

Figure 1 shows that the “End Poverty” pillar appears to be a valid latent construct with good indicators. The “Social Inclusion” factor however is much weaker in its validity, with some indicators like Education representing the latent construct quite well and others like the GINI coefficient having low factor loadings. The “Environment” factor is particularly poorly defined and incoherent. There seems to be only a weak relation between the various environment indicators.

A Principal Component Analysis (PCA) further emphasizes the inconsistencies of the SDG framework (see 2A). Three of the Environment goals – CO2 emissions, Protected Land and Protected Sea – load in the opposite direction on the first factor dimension than the poverty and social-inclusion related goals. An Exploratory Factor Analysis (EFA) then suggests a model that departs considerably from the SDG concept (see Figure 2B). A “Development” factor is extracted that represents development in classical terms, i.e. reducing poverty and fostering socio-economic inclusion. Within the Development factor, the CO2 emissions indicator has a negative loading, suggesting an inherent conflict between classic development and ecological sustainability. The other Environment indicator, Air Pollution, loads positively to the Development factor, but with a weak factor loading. All the other indicators from the Environment pillar and two from the Social Inclusion pillar (Youth Unemployment and Women Parliament) were dropped by the EFA, which suggests that they are neither related with Development nor with each other, as they do not build a factor on their own. This means they do not conflict with the other SDGs but they also do not correlate with them positively.

The EFA gives a second factor of “Inequality & Violence”, which shows that inequality measured by the GINI coefficient is related to Violence. Galtung and Hoeivik (1971) suggest that “direct violence” has roots in “structural violence”, that is in injustice, exploitation and extreme inequality that stunts disadvantaged peoples’ ability to develop their human potential. Our EFA supports this theory, and identifies Violence and inequality (GINI) as lying together on a separate dimension to the other developmental goals.

4. Modeling the Inconsistency of Sustainable Development

The factor analyses results suggest that SDGs will be difficult to attain if we continue with the development model that most countries have adopted in the past. However, the analyses do not reveal what the major problems are in
Figure 1: CFA Results. The factors are shown in ellipses, while the observed indicators in rectangles. The numbers above the arrows are standardized factor loadings. *variables were reverse coded, to ensure a common directionality of all observed indicators. **1/CO2. Model Fits: CFI: 0.91, TLI: 0.89, SRMR: 0.08, RMSEA: 0.06.
Figure 2: (A) PCA Factor Map of 16 SDG indicators with Factor Dimension 1 having a proportion of explained variance of 37.3% and Factor Dimension 2 of 11.6%. The axes depict standardized factor loadings. The color shows the contribution of each indicator to the factor solution: indicators in red contribute strongly, those in blue weakly. (B) EFA Results. The factors are shown in ellipses, while the observed indicators in rectangles. The numbers above the arrows are standardized factor loadings. *variables were reverse coded. CFA Model Fits for the EFA model: CFI: 0.96, TLI: 0.95, SRMR: 0.06, RMSEA: 0.06.
Figure 3: EFA-Biplot with Factor Dimension 1 having a proportion of explained variance of 73.5% and Factor Dimension 2 of 21.1%. The EFA suggests a single factor solution. A CFA for this factor, $L$, comprising of CO2 emissions (factor loading: -0.53), Child Mortality (factor loading: 0.82) and Education (reverse coded, factor loading: 0.98) has good Model Fits: CFI: 0.97, TLI: 0.93, RMSEA: 0.03 and SRMR: 0.06.

obtaining sustainable development. To investigate this in more detail we first chose one indicator for each of the three SDG pillars to create a latent variable, $L$, representing the inconsistent development. To represent each pillar we chose those indicators with the highest EFA/CFA loadings and best available data, namely Child Mortality, Education and CO2 emissions. We then performed a new CFA on these three variables (see 3) and used the CFA factor scores to create the latent variable $L$ (see Appendix A2 Methods).

We then applied a Feature Selection Algorithm (see Appendix A2.2) to scan through the large number of potential indicators in our dataset to find the twelve most relevant predictors for changes in the latent variable $L$ (see Appendix A3 Results, Table A1). These twelve indicators were then used to fit a dynamical system model (see Appendix A2 Methods) for changes in $L$. The best model, according to the Bayes Factor (see Appendix A3 Results, Table A2), for rate of change of $L$ is

$$0.46 \frac{D}{G} + 0.002G^3 - 0.02G^2 - 0.01DF_r - 0.06\frac{R_f}{J} - 0.002N^2_d. \tag{1}$$

The equation 1 terms reveal what factors are associated with development. The effects captured in the equation terms are moreover visualized in Figure 4. Note that a negative value of equation 1 implies a decrease in child mortality and in the proportion of people without secondary school education, but an increase in CO2 emissions. Conversely, a positive value of equation 1 has the opposite implication. The first term indicates that high debts ($D/G$ is the ratio net foreign assets, $D$, to log GDP per capita, $G$) prevent a reduction in $L$, with high indebtedness being particularly detrimental to socio-economic development if countries are poor. The second and third terms indicate that countries with larger GDP reduce $L$ faster (see Figure 4), but this effect is cancelled out as the economy reaches the size of a modern Western economy (i.e. approximately $G = 10$). This is further exacerbated when debts are high. Thus, economic growth and low indebtedness reduce poverty and facilitate socio-economic inclusion but simultaneously harm the environment. The fourth term arises because poor countries (i.e. with high indebtedness, $D$, and high fertility rates, $F_r$) have a general tendency to make progress in terms of socio-economic development, while environmental depletion tends to increase. The fifth term reflects that women’s economic rights, $R_f$, is positively associated with socio-economic development, particularly in countries with weak judicial institutions, $J$. The last term shows that $L$ decreases as an accelerating function of natural depletion costs, $N_d$.

To sift out the reasons for sustainable development inconsistencies, we then investigated the indicators associated with changes in the three SDG pillars: Child Mortality, Education and CO2 emissions (see Appendix A3 Results, Table A2). We found that decreases in Child Mortality are best predicted by

$$-0.03T_fG + 0.86M - 6.4\frac{M}{G} - 0.001F^3_r. \tag{2}$$

The first term shows that a combination of high log GDP, $G$, and trade freedom, $T_f$, is associated with a reduction
in Child Mortality. The second and third term indicate that efficient health programs such as measles immunization, $M$, reduce Child Mortality in poor countries. Once countries are rich, common diseases such as measles are mostly eradicated and therefore the positive effect of targeted health programmes ceases. The last term finally implies that poor countries that have typically high fertility rates, $F_r$, tend generally to develop and therefore to reduce Child Mortality.

Changes in Education were best predicted by

$$-0.01G - 0.03W_g^2 + 0.001CG + 0.16\frac{W_g}{C}. \tag{3}$$

The first and third term show that higher log GDP per capita, $G$, and government spending, $W_g$, predict decreases in the proportion of people without secondary school education. In the second and fourth terms, final consumption expenditure, $C$, combines with the other two indicators to signify that in rich countries, where the majority has presumably attained secondary education, the reduction ceases.

Finally, changes in CO2 emissions were best predicted by

$$0.00002\frac{Nd}{Er} - 0.0004G^3 + 0.11GE_m - 0.11CE_m + 0.004GC - 0.003\frac{C}{E_m}. \tag{4}$$

The equation combines several non-linear terms, involving natural depletion costs ($N_d$), renewable energy production ($E_r$), log GDP per capita ($G$), particulate emission damage ($E_m$) and final consumption expenditure ($C$). The model is highly complex and shows how the various factors interact in various non-linear ways. Combined, these terms show that poor countries have low CO2 emissions, that then rise with growing economy and consumption until countries have reached very high wealth levels, at which point CO2 emissions can be expected to stall, though at this stage the CO2 emissions levels of a country will be already unsustainably high. CO2 emissions are proportional to overall natural depletion costs per unit of energy produced through biomass and they are coupled with particulate emission damage, thus with detrimental effects of environmental pollution on human health.
The four equations 1, 2, 3 and 4 provide some explanation for the inconsistency of sustainable development. All models include GDP per capita, which has overall a positive effect on reducing poverty (equation 1, 2) and increasing socio-economic inclusion (equation 1, 3), but a mainly negative effect on reducing CO2 emissions (equation 1, 4). It is the current economic system that is based on growth and consumption (C in equation 3, 4) that makes some of the SDGs incompatible. As every nation seeks to increase economic growth to meet the rising standard of living expectations of its population, nature is under-prioritized (Rich, 2013; Redclift, 2010; Managi and Kaneko, 2009; Jorgenson, 2010; Pao and Tsai, 2010). But the models also reveal factors (indicators unique to 2, 3 and 4) that have beneficial effects on one goal, without having simultaneously adverse effects on other goals. These include extensive health programs for reducing Child Mortality, government spending on Education, and renewable energy production for reducing CO2 emissions.

The models do not explicitly take into account technological innovations, which some claim will ensure that climate change and other environmental problems can be addressed in future (Grubb, 2013; Ridley, 2010; OECD (Organisation for Economic Co-Operation and Development, 2011). The feature selection algorithm and model selection procedure did not pick specific technological indicators from the vast dataset as good predictors, except for the renewable energy production ($E_r$) indicator. In fact the fifth term in equation 4 can be interpreted as capturing technological innovation targeted at reducing environmental degradation, i.e. climate change. Capturing overall technological change and its potential (future) contribution to mitigating environmental depletions is however extremely challenging. Arrow et al. (2012) for instance use time itself in their models on sustainable development to account for effects of time-varying factors such as technological change. However, this approach was widely criticised as inappropriate (Solow, 2012). We therefore restrained from including time in our models, but the models can be easily extended to include time as a predictor.

5. Monitoring Sustainable Development

Giving the current socio-economic development “is stacked against nature” as Dasgupta (2013) noted, it will be necessary to develop new development models and agendas; and along with them new measures of sustainable development, progress and nation’s wealth that will help to monitor worldwide sustainable development advances.

One of the most elaborated and comprehensive recent approaches to measure sustainable development and nations’ wealth is the Inclusive Wealth Index. Dasgupta (2010) defines sustainable development in terms of intergenerational wellbeing, which Dasgupta tries to capture with the inclusive or comprehensive wealth measure (Dasgupta, 2007a; Arrow et al., 2012). This measure represents a society’s stock of all its capital assets (reproducible/productive capital, human capital and natural capital) and their changes over time accounting for population growth and technological change. These various capital assets form a society’s productive base which is a means to protecting and promoting well-being across the generations. Therefore, economic development is only sustainable, if the change in inclusive wealth over time is positive and is likely to increase in future (ibid.). This is usually the case if consumption per capita does not exceed net domestic product per capita, which is interpreted as GDP per capita minus the depreciation of those capital assets and/or if (government) investment (e.g. in health, preserving ecosystems) is positive (Dasgupta, 2013). In consequence, sustainable development can involve excessively high rates of (government) investment, which can be a burden on the current generation (ibid.). Dasgupta shows empirically that the Inclusive Wealth Index frequently indicates a decline in intergenerational wellbeing while GDP per capita and HDI increased (Dasgupta, 2007a,b, 2010, 2013), demonstrating the unsuitability of these traditional indices for measuring sustainable development. However, Dasgupta’s theoretically elegant approach is currently severely limited by cross-country, time-series data availability issues, particularly when it comes to environmental data (Arrow et al., 2012; Dasgupta, 2013). Moreover, the approach would accept a temporary conflict between different capital assets and thus between socio-economic and environmental development goals as long as the overall sum of the inclusive wealth change remains positive and is expected to grow in future. The expected future increase is however difficult to estimate accurately, particular as the model does not take into account nonlinear development dynamics (Spaiser et al., 2014; Ranganathan et al., 2015b).

We suggest here an alternative, data-driven approach for measuring and monitoring sustainable development, accounting for data availability issues and nonlinear dynamics in development processes. First of all, we can use models for single SDG goals, equations 2, 3 and 4, to monitor progress of the SDGs separately. These models can be moreover used to make future predictions about likely development trajectories and potential development traps (Spaiser et al., 2014; Ranganathan et al., 2015a). For an overall index to monitor sustainable development, we can use
the rate of change in our latent variable. The model in equation 1 is used to calculate scores for each country and year, using country- and year-specific initial values for the predictors in the model (SDG index 1). As such the resulting index scores represent the expected changes in a country’s development status and thus the expected progress of a country.

Alternatively, an overall index can be constructed directly from the models for the single SDG goals using a Bayesian model combination approach (see Appendix A2.3). Such an approach would result in an alternative model for the latent variable:

$$\frac{dL}{dt} = 0.01GC - 0.07G + 2.1 \frac{W_g}{C} - 0.46W_g^2 - 2e^{-0.4F_r} + 7.1e^{-0.06N_d}$$

In this model too negative values imply a decrease in child mortality and in the proportion of people without secondary school education, but an increase in CO2 emissions. Conversely, a positive value of equation 5 has the opposite implication. Model term one and two show that higher log GDP per capita ($G$) is generally beneficial for socio-economic development until rich countries reach very high GDP levels (along with high consumption expenditure ($C$) levels) at which point the positive effect ceases. The third and fourth terms reveal that government investment ($W_g$) is important in lowering poverty and increasing socio-economic inclusion, once the government spending is high enough. However, in rich countries (high levels of final consumption expenditure $C$), further improvements in socio-economic development through government spending cannot be achieved. However, government spending may then contribute to the mitigation of high CO2 emissions. The fifth term indicates that poorer countries (high fertility levels ($F_r$)), have a general tendency to develop. The sixth term finally shows that the socio-economic development comes at natural depletion costs ($N_d$), which can be reduced however if renewable energy ($E_r$) production is increased. This alternative model can be used as an alternative overall index to monitor Sustainable Development Goals (SDG index 2).

Figure 6 shows two world maps with countries coloured based on their SDG index scores for the year 2011. We can see that the two indices do not make always the same predictions about countries’ development. The first SDG
Table 1: Predictive Power of the two SDG indices, SDG index 1 is based on model 1 and SDG index 2 is based on model 5 in comparison to HDI and GDP per capita, based on $R^2$

| SDG index 1 predicts... | SDG index 2 predicts... | HDI predicts... | GDP per capita predicts... |
|-------------------------|-------------------------|-----------------|--------------------------|
| 54% of changes in child mortality | 46% of changes in child mortality | 41% of changes in child mortality | 17% of changes in child mortality |
| 6% of changes in education | 12% of changes in education | 4% of changes in education | 2% of changes in education |
| 21% of changes in CO2 emissions | 48% of changes in CO2 emissions | 0.7% of changes in CO2 emissions | 0.4% of changes in CO2 emissions |
| 16% of changes in $L$ | 14% of changes in $L$ | 7% of changes in $L$ | 4% of changes in $L$ |

index, based on equation 1, suggests that most counties in the East, such as Russia, China, Kazakhstan, India, as well as some South American (e.g. Mexico), Arabic (e.g. Tunisia) and African (e.g. Mauritania) countries are performing well in reducing poverty and increasing socio-economic inclusion. However, given the incompatibility of the SDGs, this positive development comes along with increasing CO2 emissions. Rich developed countries such as the US, Canada, Australia and UK have reached already high levels of socio-economic development, thus there is little room for further development. The SDG index rather suggests that poverty and socio-economic exclusion may be rising again in those countries. On the other hand, we would expect that the CO2 emissions will level off or even slightly decrease in those countries in the long run, though on a high unsustainable level. Some South American (e.g Brazil), Asian (e.g Thailand) and African (e.g. South Africa) countries on the other hand seem to be stuck with no or only very slow socio-economic development or even some drawback in terms of socio-economic development.

The second SDG index based on equation 5 suggests similar results for rich, developed countries. But according to the second SDG index Russia seems to be stuck with no or only very slow socio-economic development. China, India and Kazakhstan on the other hand are again performing rather well in terms of socio-economic development. However, their socio-economic development goes hand in hand with rising CO2 emissions. Moreover, the second index suggests that several African countries (e.g. Angola, Kenya) are not stalling or regressing, as suggested by the first SDG index, but in fact making some progress in terms of socio-economic development.

The differences in the predictions for various countries result from the different conceptualizations of the two models in equation 1 and 5. The first model assumes a true latent variable with the three components being only some observable indicators for this latent phenomena, which ultimately goes beyond those three indicators. Thus, the model seeks to predict changes in this latent variable primarily and not changes in its components. On the other hand, model 5 is much closer to the three components, given it was built from the three original component models. Thus, model 5 would tend to primarily predict changes in the three components and to a lesser extent changes in the true latent variable $L$.

To decide which SDG index is more suitable for monitoring sustainable development and to compare the performance of our predictive SDG indices with widely used indices of economic development, GDP per capita and HDI, we compared how well these indices are doing in predicting changes in the latent variable $L$ and in its three components, Child Mortality, Education and CO2 emissions. We thus fitted models where the indices would predict changes in the three $L$ components and in $L$ itself and compared the Bayes factors of these models. We did not include a comparison with the Inclusive Wealth Index or other recently suggested development indices such as the Happy Planet Index (Abdallah et al., 2012) or Social Progress Index (Porter et al., 2013), because these indices are currently available only for two or three years (and rather recent years) and do not always cover the majority of worlds’ countries. Figure 7 and Table 1 summarise the evaluation results.

Our SDG indices perform better than the common indices for economic development HDI and GDP per capita. Moreover, the SDG index 1 based on equation 1 seems to be a better predictor for the overall SDG process and for Child Mortality comparing to the SDG index 2 based on model 5. But, it performs less well in predicting changes in education and CO2 emissions comparing to the second SDG index. It should be noted, that the inherent contradiction between development and sustainability is encoded in equations 1 and 5 and therefore in the SDG indices. They do not resolve the conflict and remain therefore dialectic.
Figure 6: World maps with countries being coloured based on their SDG indices scores in 2011. White colour indicates that an index could not be calculated due to missing data in one or several of the predictors in the two $dL$ models. (a) shows the index colouring based on equation 1 and (b) shows the index colouring based on equation 5.
Figure 7: Bayes Factors for models predicting changes in L, Child Mortality, Education and CO2 emissions with either SDG index1, SDG index2, HDI or GDP per capita as predictors
6. Conclusion

Over the past 30 years there has been an inherent and apparently unavoidable conflict between socio-economic development and ecological sustainability (Rich, 2013; Redclift, 2002, 2010; Jorgenson, 2010). Our results have quantified this inconsistency and showed that economic growth fulfills socio-economic goals while simultaneously hindering environmental goals. On a more positive note, our models identify factors, which can contribute to socio-economic development (health programs, government spending) on the one hand and ecological sustainability (renewable energy) on the other, without triggering the conflict between incompatible SDGs. The decoupling of the two SDG indicators for Violence and Inequality (see Figure 2) offers the possibility that improvement can be made on these goals independent of the other 14 goals.

While our data analyses reveal which indicators are associated with improvements towards the SDGs, they do not disclose which actions will allow us to achieve the SDGs. Data analyses show the world as it is, not as it could be. (Field-)experimental approaches, as for instance developed by the MIT Poverty Action Lab (Tollefson, 2015; Schill et al., 2015), could contribute to testing “alternative worlds”. Similarly, case studies (Duit, 2014; Guerrero et al., 2016; Crépin, 2003; Dearing et al., 2014), e.g. of the Bolivian case, the first country to grant nature equal rights to humans in its constitution (Bolivian Democratic Government, 2010) could provide insights into alternative development trajectories. Data analyses are however useful in monitoring changes. As more reliable development data becomes available along with the UN “data revolution for sustainable development” (United National Department of Economic and Social Affairs, 2015), we can both monitor progress towards the SDGs and, by revising our models and indices in light of improved data, develop a deeper understanding of the path which our planet is taking.

We acknowledge that sustainable development is a long-term process with potentially unforeseen future turns like major technological innovations for instance, not present in data records. On the other hand, technological innovations follow incentives and since the environment is constantly underpriced and disregarded in the present economic system, the innovators have little reason to develop technologies that focus on reducing environmental depreciation (Dasgupta and Ehrlich, 2013). Moreover, historical evidence suggests that successful technological response cannot be guaranteed (ibid.). We thus have to try and find responses to the incompatibility of economic development and environmental sustainability within our present capabilities.

Appendix A1 Data

The comprehensive World Bank API dataset of 1364 indicators covers indicators on agriculture, trade and investment, imports and exports, economic performance, IT infrastructure, remittances, capital and financial sector, debts and savings, development aid, water pollution, energy sector (including sustainable energy production), greenhouse gases and other pollutants emissions, threatened species (though the data is practically not usable due to incompleteness), population and demography, sanitation, terrestrial and marine protected area, taxes, government expenses, household and national consumption expenditure, business activities and regulations, innovation (patent application etc.), transport and transport infrastructure, industry, education (including literacy rate), gender equality, health and health infrastructure, nutrition, poverty, inequality, labour and employment, migration and refugees, urbanization, tourism and finally governance (including corruption, rule of law, etc.) (World Bank Development Data Group, 2010, 2014; Kaufmann et al., 2010).

The data quality varies considerably with some indicators such as GDP having almost complete data coverage, and others having hardly any data coverage at all (e.g. environmental indicators but also some poverty indicators) at least if one is interested in dynamical (time series) analysis. Generally it is quite evident that while economic data is largely complete, data on environment (with exception of CO2 emissions), equality, social matters etc. suffers from incompleteness and bad quality. This alone creates a misbalance and bias for all analyses involving this data.

As noted in the main text, to this World Bank dataset we have added datasets from Polity IV containing measures on democracy vs. autocracy and regime durability (Marshall et al., 2014), CIRI Human Rights Data Project containing various human rights indices, such as the physical integrity rights index consisting of measures that count the cases of disappearance, extra-judicial killing, political imprisonment and torture in a country (Cingranelli et al., 2013), Centre for Systemic Peace data containing data on conflicts (interstate wars, civil wars, etc.), state fragility, governance (effectiveness and legitimacy in security, politics, economy and social matters), (compiled by Monty G. Marshall, 2014), Forcibly Displaces Populations Data provided by the U.S. Committee for Refugees and Immigration (USCRI)
and Internal Displacement Monitoring Centre (IDMC) and compiled by the Centre for Systemic Peace (Marshall and Cole, 2014), the Freedom House data containing the indices for political and civil rights as well as freedom of press (Freedom House, 2014a,b) and the Heritage Foundation/Wall Street Journal data containing measures on economic freedom, such as property rights, freedom from corruption, fiscal freedom, government spending, business freedom, labor freedom, trade freedom, investment freedom and financial freedom (Miller et al., 2014). The quality of these datasets is usually considerably good, given dedicated research groups work on producing and maintaining these datasets, but the time coverage varies significantly. The combined dataset (including the World Bank dataset) consists of 1423 economic, social, political and environmental indicators for 217 countries covering the years 1980 to 2014.

One of the major challenges to a data-driven analysis of the SDGs is that the currently available data is still quite incomplete and often of bad quality and reliability as mentioned a above. While we started with 1423 indicators, we finally could use only 233. Partly the reduction is due to conceptual duplicates, for instance, GDP per capita is given multiple times, as GDP per capita can be measured in current and constant 2005 US$, in current and constant 2011 international $, in current and constant local currency, etc. We removed all duplicates, aiming at using consistent measures. For instance, we always used current US$ for monetary measures. But quite often we had to remove indices because they had too many missing values, e.g. 80% even after interpolation, which would inhibit any reasonable modelling attempts.

The data was linearly interpolated to obtain as many data points as possible for reasonable analysis. We also rescaled a range of measures to align the value ranges of different variables better to each other and to produce relative measures by adjusting total measures (e.g. CO2 emissions) by the population size of a country (e.g. that way obtaining CO2 emissions per capita). Monetary measures like GDP, GNI etc. were moreover transformed into log scales to obtain value ranges that are better aligned with other variable scales.

A1.1 SDG indicators and reverse coding

From the 233 variables that we could use for our analyses 16 were representing the 17 SDGs (see main manuscript). The 16 variables were suggested by the UN as measures for the 17 SDGs and its 169 sub-targets (Leadership Council of the Sustainable Development Solutions Network, 2015; United National Department of Economic and Social Affairs, 2015). Where multiple indicators were listed by the UN as potential measures for the 16 SDGs and their sub-targets, we selected those with the best data coverage. Since the 17th goal is an administrative goal, international partnership for the SDGs, and therefore not meant to be measured or quantified, it was not further included in the analyses. The SDGs are often reported to consist of four pillars, besides the three mentioned in the paper (poverty reduction, socio-economic inclusion and ecological sustainability), the fourth pillar is focusing on peace, government efficacy and rule of law. This pillar is rather marginal in the overall SDG agenda, given it is represented only by one of the 17 goals, namely goal 16. One indicator representing this goal, homicide rate per 1000 inhabitant ("Violence", Vio), was nevertheless included in the factor analysis.

The 16 indicators representing the 16 SDG goals are: "GINI" coefficient measuring inequality, "Violence" (Vio), number of homicides per 1000 inhabitants, "Women Parliament" (WP), percentage of women in national parliaments, reverse coded this variable measure the percentage of men in national parliaments, "Youth Unemployment" (YU), percentage of people under 25 without employment, "Education" (Ed), percentage of children getting secondary school education, reverse coded this variable measures the percentage of children excluded from secondary school education, "Child Mortality" (CM), number of children under five dying per 1000 births, "Poverty" (Pov), percentage of people who have less than 1.25$ per day, "Hunger" (Hun), percentage of people below minimum level of dietary energy intake, "Water" (Wat), percentage of people having access to drinking water, reverse coded this variable measures the percentage of people not having access to drinking water, "Sanitation" (San), percentage of people having access to sanitation facilities, reverse coded this variable measures percentage of people not having access to sanitation facilities, "Internet" (Int), number of people having access to the Internet per 1000 inhabitants, reverse coded this variables measures the number of people per 1000 inhabitants who do not have access to the Internet, "CO2", CO2 emissions per capita, "Protected Land" (PL), percentage of land protected by law, reverse coded this variable measures the percentage of land not protected by law, "Protected Sea" (PS), percentage of bordering sea area protected by law, reverse coded this variables measures percentage of bordering sea area not protected by law, "Alternative Energy" (AE), percentage of energy production coming from alternative, non-fossil energy sources, reverse coded this variable measures percentage of energy production coming from conventional, fossil energy sources and "Air Pollution" (AP), number of deaths due to air pollution per 1000 inhabitants. Some variables were reverse coded to ensure a common...
directionality of all observed indicators, with low values being desirable in terms of the SDG agenda (e.g. decrease child mortality, decrease poverty, decrease the share of the population without secondary school education, decrease environmental pollution, etc.).

Appendix A2 Methodology

A2.1 Factor Analysis

As a first step in our analysis, we investigated the internal consistency of the Sustainable Development Goals (SDGs). We conducted a Confirmatory Factor Analyses (CFA), a special case of a covariance-based structural equation model (Harrington, 2008; Reinecke, 2005). The CFA assumes a theoretical or conceptual model with defined observed variables, which serve as indicators for the postulated latent factor or factors. The observed variables build together with measurement errors a linear combination of the theoretical construct, i.e. the latent factor. The estimation is realised with a Robust Maximum Likelihood approach. Model indices such as the Root Mean Square Error of Approximation (RMSEA), the Standardised Rood Mean Square Residual (SRMR), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI) are used to assess the overall goodness of the CFA model.

Standardized factor loadings and z-scores, which describe the relation between the estimated value (factor loading) and its standard error, are used to evaluate the contribution of the observation variables to the latent factor. Standardized factor loadings with values closer to -1 or +1 indicate a strong predictability of the observed indicators by the factor. The CFA can be hierarchical involving first and second order latent constructs. The observed variables construct two or more first order latent factors and these first order latent factors then construct one or several overarching second order factors (Harrington, 2008; Reinecke, 2005). With the CFA approach we thus tested the conceptual SDG framework against available data.1

As the results (see main manuscript) suggested that the SDG framework is conceptually inconsistent with the data we also conducted an Exploratory Factor Analysis (EFA) to examine what latent structure is suggested by the data (Thompson, 2004). The EFA is used to test whether the correlation structure of the observed variables allows the extraction of one or several factors, and thus, to what extent the observed variables can be predicted by one or several latent constructs. Generally two approaches can be used for extracting latent factors, the Principal Component Analysis (PCA) or the Principal Axis Analysis (PAF). The two approaches are quite distinct in their underlying assumptions and interpretations, but they often produce the same results.

The PAF focuses on shared variance, not on sources of error that are considered to be unique to the single observed measurements. The PCA, on the other hand, focuses on explaining all of the variance of the observed indicators through a set of latent factors (Thompson, 2004). We used the PCA approach combined with the Varimax rotation method, which seeks to extract a simple factor solution where each factor has a small number of large factor loading and a large number of zero (or small) factor loadings. This simplifies the interpretation of the factor analysis results, because each observed indicator tends to be associated with one (or a small number) of factors and each factor represents only a minimal subset of observed indicators. We also verified our results with the PAF approach.2 To test the EFA results we finalized the factor analyses by conducting a CFA with the EFA-suggested model.

Based on the results of the factor analyses and under considerations of data availability we selected three indicators (Child Mortality, Education, reverse coded, and CO2 emissions per capita), each representing one of the three core SDG pillars (reducing poverty, socio-economic inclusion, ecological sustainability). We then conducted another CFA to see if the three indicators selected were suitable to represent a latent SDG variable. We used the CFA factor scores from this procedure to create the latent variable L for subsequent analyses. In particular, we used regression scores. This is a refined method for obtaining factor scores, where the resulting scores are linear combinations of the observed variables based on the shared variance between the indicator (item) and the factor (latent variable) and the variance that is not measured (Christine DiStefano and Mindrila, 2009). Standardized information is used to produce standardized factor scores similar to a z-score metric, where values range from approximately -3.0 to +3.0. Generally, refined

1We used the R package sem to estimate CFAs for the SDGs (Fox, 2006).
2We used the R function factanal, which is included in the R package stats, one of the essential R packages developed by the R Core team, to estimate EFAs for SDGs, moreover the R function factor.pa, implemented in the R package psych (Revelle, 2015) and the package FactoMineR (Husson et al., 2015) were used to verify and visualise factor analysis results.
methods of factor score calculation “aim to maximize validity by producing factor scores that are highly correlated with a given factor and to obtain unbiased estimates of the true factor scores” (ibid.). Regression scores specifically predict the location of each individual on the factor and take into account the correlation between the factors and observed indicators (through item loadings), but also the correlation among the observed indicators (ibid.)

A2.2 Feature Selection Algorithm

The latent variable \( L \) represents our main dependent variable. In the next step our goal was to identify relevant indicators (not including the 16 SDG indicators used in the factor analyses) that would predict within a mathematical model how progress on the SDGs can be made. Removing the 16 SDG indicators from our dataset of 233 indicators, we were still left with 217 potential predictors; too many for reasonable modeling. We therefore used a Feature Selection Algorithm (Guyon and Elisseeff, 2003; Mehmood et al., 2012) to preselect the most relevant predictors. We used specifically the Variable Elimination Algorithm (Mehmood et al., 2012), a supervised feature selection machine learning method based on partial least square regression.

The uninformative variable elimination is a wrapper method, based on procedures iterating between model fitting and variable selection. The search algorithm extracts a subset of relevant variables and evaluates each subset by fitting a model (ibid.). Monte-Carlo-based uninformative variable elimination methods that use randomised search in the selection of variables subsets (Mehmood et al., 2012; Centner et al., 1996), have proved to be successful techniques in selecting relevant predictors (Cai et al., 2008; Han et al., 2008). The relevant variables are those that have both large mean values and small standard deviations with respect to the regression coefficient, thus the measure of variable importance is constructed as follows:

\[
c = \frac{\text{means}(s)}{\text{sd}(s)}
\]

where \( s \) is the coefficient vector for the \( i \)-th variable, generated by Monte-Carlo resampling. The variables can be ranked based on this measure, with the most relevant ones having the highest \( c \) value (Centner et al., 1996; Xiao et al., 2014). The dependent variables at this stage as in later modeling stages were changes in the latent variables \( L \) as well as changes in the three item components of the latent variable, thus changes in child mortality (\( CM \)), secondary school education, reverse coded (\( Ed \)) and CO2 emissions per capita (\( CO_2 \)). We therefore pre-selected four sets of most relevant variables, respectively for the changes in \( L \), \( CM \), \( Ed \), and \( CO_2 \).

A2.3 Dynamical Systems Modeling

The pre-selected variables entered at the next stage a dynamical system modeling procedure, as outlined in Ranganathan et al. (2014a). We used the data to inform model selection from a pool of potential models by fitting polynomial differential equations that are capable of capturing various linear and non-linear dynamical patterns in the data. Changes in \( L \) between times \( t \) and \( t + 1 \) were modelled as a function of all possible combinations of polynomial terms involving preselected predictors at time \( t \) with varying complexity (number of polynomial terms included in the model).

The procedure in Ranganathan et al. (2014a) was however modified to avoid the limitation of modeling with maximum four predictors. The limitation was originally set because the number of possible models increases exponentially with each variable added. Having two indicators and a set of 17 polynomial terms (e.g. \( x, x^2, x^3, \frac{x}{y}, xy, \frac{1}{y} \), etc., see (Ranganathan et al., 2014a)) with a model complexity of maximum 6 polynomial terms, results already in 12,376 possible models, increasing the number of indicators to four variables would result in 97 possible polynomial terms, if keeping the type of terms defined for the two-indicator case and this would result in over 3,464,840 possible models, even if the model complexity is limited to a maximum of 4 polynomial terms.\(^4\)

By iterating through various combinations of indicators sets and polynomial term sets, eliminating all those variables and polynomial terms with bad model fittings in several combination rounds and keeping only the relevant terms and variables for the next combination round where other variables and thus terms would be added, we could include up to 20 predictors. Also, in contrast to the approach outlined in (Ranganathan et al., 2014a) we only consider models

\(^3\)We used the R package \textit{enpls} to extract the relevant variables for our SDGs analysis (Xiao et al., 2014).

\(^4\)We used our own R package \textit{bdynsys} for the dynamical systems modeling approach and computation of Bayesian factors (Ranganathan et al., 2014b), however extended the package functions as described above.
for changes in the dependent variables. Models that predict changes in the predictors were not elaborated, as we were only interested in identifying predictive models for the SDGs. We thus distinguished strictly between dependent (SDGs) and independent variables (predictors). Consequently, the dependent variables were not allowed to appear in the models as "self-predictors", as was the case in the original methodology (Ranganathan et al., 2014a). The dynamical systems modeling approach was applied separately to \( L \) and its three components \( CM, Ed \) and \( CO2 \). Bayesian model selection as outlined in (Ranganathan et al., 2014a) was used to identify the overall best-fitting model.

As a result we had a dynamical systems model with four differential equations as an output, one for changes in \( L \), one for changes in \( CM \), one for changes in \( Ed \) and one for changes in \( CO2 \). In a follow-up analyses we also used a Bayesian Model Combination approach (Kristine Monteith and James L. Carroll and Kevin Seppi and Tony Martinez, 2011) to derive an alternative mathematical model for \( L \). We took the three models for the three components (\( CM, Ed, CO2 \)) and fitted all possible combinations of them, picking finally the one with the highest Bayesian factor. As such this model represents another model variant predicting changes in the latent variable \( L \), however closer to the three components that construct the latent variable.

Appendix A3 Additional Results and Analyses

A3.1 Factor Analysis Results

The consistency of the SDGs was first tested using a Confirmatory Factor Analysis (CFA). In the specification of the CFA we had to take \( CO2 \) to achieve somewhat reasonable factor loadings and model fits, even though this step suggests inconsistency in the directionality of the SDGs, i.e. a conflict between mitigation of \( CO2 \) emissions and socio-economic development. A CFA model with unmodified \( CO2 \) emissions indicator yielded an even poorer model (CFA Model Fits: CFI: 0.84, TLI: 0.81, SRMR: 0.13, RMSEA: 0.08) than the one presented in the main manuscript (CFA Model Fits: CFI: 0.91, TLI: 0.89, SRMR: 0.08, RMSEA: 0.06). As already pointed out in the main manuscript there seems to be only a weak relation between the various environmental indicators. This may be due to poor data availability for environment indicators, particular the data for Protected Land and Protected Sea was very scarce.

Given the CFA revealed a weak conceptualisation of the SDG agenda, we performed subsequently an Exploratory Factor Analysis (EFA) as described in the main manuscript. The EFA suggested a two factor solution with the factor 'Development' that included all the poverty-related indicators (Poverty, Child Mortality, Hunger, Water, and Sanitation), two socio-economic inclusion related indicators (Education and Internet) and two environment related indicators (\( CO2 \) emissions and Air Pollution), though \( CO2 \) (please note, here we used \( CO2 \) not \( CO2^2 \)!) has a negative factor loading, which indicates the inherent conflict between the environmental and socio-economic dimension of development. The second factor, that is not correlated significantly with the factor 'Development', is the factor 'Inequality & Violence', consisting of the two indicators Violence and GINI. All remaining indicators, Youth Unemployment, Women Parliament, Protected Land, Protected Sea and Alternative Energy did not not load significantly to any of the two factors 'Development' or 'Inequality & Violence' and did not build a factor on their own. These indicators are therefore completely independent of each other and of the SDG indicators included in the two factor solution. The test of the hypothesis that two factors are sufficient was significant with \( X^2(df = 89) = 1675.46 \) with \( p < 0.01 \) and the Tucker Lewis Index of factoring reliability was 0.78. Furthermore, Cronbachs Alpha for the first factor (Development) was sufficiently high with 0.65, while the Inter-Item correlation for second factor (Inequality & Violence) was 0.34.

We then performed a CFA for the two factors 'Development' and 'Inequality & Violence' and their indicators suggested by the EFA, excluding the unrelated variables (Youth Unemployment, Women Parliament, Protected Land, Protected Sea and Alternative Energy) from the CFA model. The CFA model had good Model Fits with CFI: 0.96, TLI: 0.95, SRMR: 0.06, RMSEA: 0.06.

A3.2 Feature Selection Results

To identify the most relevant predictors from a set of 217 potential predictors we used a feature selection algorithm (see S2.2) and selected for each dependent variable the twelve highest scoring and therefore most relevant predictors. Table A1 shows the selected predictors with their respective variable importance scores.

For all the dependent variables we can see a mix of economic, social and political indicators that were selected as predictors. For changes in child mortality, health-related indicators such as health expenditure \( (H) \) and the implementation of immunization programs for new borns, such as the measles immunization \( (M) \), are of big importance.
Table A1: Selected predictors with their respective variable importance scores. Predictors identified with the variable elimination algorithm based on partial least square regression.

| Child Mortality (dCM) | Education (dEd) | CO2 emissions (dCO2) | Latent Variable (dL) |
|-----------------------|-----------------|----------------------|----------------------|
| Health expenditure per capita (% of GDP) (H): 4.69 | GDP per capita (G): 2.16 | Household final consumption expenditure (C_h): 2.11 | Net foreign assets (undeberedness) (D): 1.39 |
| Immunization, measles (% of children ages 12-23 months) (M): 4.15 | Political rights (R): 1.30 | Final consumption expenditure (C): 1.24 | Age dependency ratio (A_d): 1.32 |
| GDP per capita (G): 4.02 | Health expenditure (% of GDP) (H): 2.09 | Control of Corruption (C): 1.19 | GDP per capita (G): 1.32 |
| Food imports (F_i): 3.86 | Industry, value added per worker (I_v): 1.56 | GDP per capita (G): 1.08 | Fertility rate (F_r): 1.11 |
| Renewable internal fresh water resources (W_f): 3.64 | Fertility rate (F_r): 1.43 | Gross fixed capital formation (K): 0.99 | Independence of Judiciary System (J): 1.03 |
| Trade Freedom (T_f): 3.24 | GDP per person employed (G_p): 1.32 | Political rights (R): 0.90 | Adjusted savings: natural resources depletion (N_r): 0.99 |
| Voice and Accountability (V): 3.03 | Energy use (E_u): 1.11 | Adjusted savings: particulate emission damage (E_p): 0.84 | Women’s economic rights (R_w): 0.99 |
| Net Adjusted Savings, net forest depletion (F_d): 2.95 | Control of Corruption (C): 1.09 | Adjusted savings: natural resources depletion (N_r): 0.78 | Religious freedom (R_f): 0.97 |
| Political Stability and Absence of Violence (P): 2.91 | Mobile phone subscriptions (M_p): 1.08 | Voice and Accountability (V): 0.70 | Mobile phone subscriptions (M_p): 0.96 |
| Agriculture raw material exports (A_r): 2.84 | Rule of Law (R_l): 1.03 | Combustible renewables and waste (E_r): 0.65 | Adjusted savings: net forest depletion (F_d): 0.93 |
| Worker’s rights (W_r): 2.80 | Final consumption expenditure (C): 0.99 | Government Spending (W_g): 0.59 | Gross national expenditure (G_e): 0.91 |
| Fertility rate (F_r): 2.78 | Government Spending (% of GDP) (W_g): 0.88 | Agriculture value added per worker (A_v): 0.57 | Health expenditure per capita (% of GDP) (H): 0.87 |

From earlier studies (Ranganathan et al., 2015b) we know that economic development contributes decisively to a decrease in child mortality. This finding is confirmed here by predictors such as GDP per capita (G) and trade freedom (T_f). Trade freedom is one of the indices of economic freedom provided by The Heritage Foundation and Wall Street Journal. It is a “composite measure of the absence of tariff and non-tariff barriers that affect imports and exports of goods and services” (Miller et al., 2014). Furthermore, two agricultural economy variables were picked as important predictors, food imports (F_i) on the one hand and agriculture raw material exports (A_r) on the other. The variable importance score does not reveal the direction of the relation between the predictor and the dependent variable. But, from the models we obtain in later analyses steps (see Table S2) we know that food imports have a negative impact on reducing child mortality. The reason for this is presumably that countries, which rely heavily on food imports are more likely to suffer from food price fluctuations, which may threaten the nutrition of poor segments of the population (Headey and Fan, 2008; Brinkman et al., 2010). Agriculture raw material exports on the other hand measures to what extent a country is agriculturally productive, which indirectly indicates whether a country is capable of meeting the nutritional needs of its population through national agricultural production.

Political indicators that are important predictors for changes in child mortality are voice and accountability (V), a Worldwide Governance Indicator (WGI) that measures "perceptions of the extent to which a country’s citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media” (Kaufmann et al., 2010), furthermore political stability and absence of violence (P), another WGI indicator that captures people’s "perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism” (ibid.), and finally worker’s rights (W_r), which is an index provided by the CIRI Human Rights Data Project (Cingranelli et al., 2013). Worker’s rights measure to what extent "the rights to freedom of association, collecting bargaining, a minimum age for the employment of children, acceptable conditions of work and protection from forced labor” are implemented (ibid.). Fertility rate (F_r) is an important social predictor for changes in child mortality as well.

Finally, two environment related indicators were suggested as important predictors for changes in child mortality, renewable internal fresh water resources (W_f), which provides an indirect measure for people’s access to drinking
water and net adjusted savings, net forest depletion \((F_d)\), which suggests that natural resource depletion and especially deforestation may affect human health and therefore newborn mortality, as has been shown in earlier studies (Walsh et al., 1993; Gottwalt, 2013; Wilcox and Ellis, 2006).

For changes in education in terms of secondary school enrolment, the feature selection algorithm suggests that particularly economic predictors are important. In a developing, prospering economies people are more likely to be better educated. Economic predictors include GDP per capita \((G)\) and GDP per person employed \((G_p)\), which measures the overall economic productivity of a country, final consumption expenditure \((C)\), which is the sum of household, private and government expenditures and as such an indicator of a country’s overall wealth and consumption level and industry, value added \((I)\), which is an indicator for productivity in the industrial sector (World Bank Development Data Group, 2010, 2014). Moreover, it is suggested that energy use \((E_g)\) is a good predictor. The variable \(E_g\) serves probably as a proxy for wealth and development of a country.

In terms of political indicators, our analyses reveal that control of corruption \((C_c)\), rule of law \((R_l)\) and governance spending \((W_g)\) are good predictors for changes in education. Control of corruption and rule of law are both WGI indicators. Control of corruption captures the "perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as ‘capture’ of the state by elites and private interests” and rule of law measures the "perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence" (Kaufmann et al., 2010). Apparently, education levels are likely to rise particularly in well-functioning and economically thriving countries, that presumably also invest in education, as the indicator governance spending suggests. Government spending is one of the indices provided by The Heritage Foundation and the Wall Street Journal (Miller et al., 2014).

Included in the list of best predictors is also the variable Mobile phone subscriptions \((M_p)\) per 1000 inhabitants, which measures the technological advancement level of a country. As a demographic measure, Fertility rate \((F_r)\) was selected by the algorithm. Fertility rate can be also seen as a detector variable to differentiate between richer and poorer countries, with poorer countries having typically higher fertility rates. Finally, and probably less self-evident, health expenditure \((H)\) is also suggested as an important indicator for education changes. This indicator may be seen as a proxy for people’s wellbeing in terms of having access to and being able to invest in their health.

The selected indicators for changes in CO2 emissions per capita are mostly from the economic domain. Former studies have shown that CO2 emissions per capita are particularly high in economically prospering and fast developing countries (Raupach et al., 2007; Ranganathan and Swain, 2014). The economic indicators that are best suited to make predictions about changes in CO2 emissions include household final consumption expenditure \((C_h)\), which is the market value of all good and services, including durable products, purchased by household (World Bank Development Data Group, 2010, 2014), thus it is an indicator of a country’s citizens’ wealth, final consumption expenditure \((C)\), which is the sum of household, private and government expenditures (ibid.), GDP per capita \((G)\), gross fixed capital formation \((K)\), which is a measure of gross domestic investment and includes land improvements, plan, machinery and equipment purchases as well as the construction of roads, railways, schools, hospitals, private residual dwellings and commercial and industrial buildings (ibid.) and finally agriculture value added per worker \((A_v)\), which is an indicator for the intensity of agricultural economy in a country. It is well established that highly industrialized agriculture (particular involving livestock and/or food production for livestock) contributes decisively to growing CO2 emissions (Anthony J. McMichael et al., 2007; Koneswaran and Nierenberg, 2008).

Political indicators too have to be taken into account when it comes to changes in CO2 emission levels, in particular political rights \((R)\), control of corruption \((C_c)\), voice and accountability \((V)\) and government spending \((W_g)\) were selected as important predictors. The political rights index is an indicator provided by Freedom House and measures the democracy level of the electoral process, to what extent political pluralism and participation are realised and the functioning of governmental bodies (Freedom House, 2014a,b). Together with voice and accountability, thus the "perceptions of the extent to which a country’s citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media” (Kaufmann et al., 2010), it measures the extent of democratization of a society. The level of democracy is apparently an important predictor for changes in CO2 emissions. Additionally, control of corruption, which measures the "perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as ‘capture’ of the state by elites and private interests” (Kaufmann et al., 2010) was selected as an important predictor. Based on former studies (Welsch, 2004; Hassaballa, 2015), we may interpret this in terms of a society’s ability to protect and preserve nature.
by fighting corruption that may contribute to circumvention of environmental standards and environment protection laws. Government spending finally may be seen as an indicators for government investment in renewable energy or natural capital recovery. Ultimately, the direction of the effect depends on the type of government investment.

Other predictors that were extracted as important features by the algorithm are directly related to the environment: Adjusted savings: particulate emission damage \( (E_p) \), Adjusted savings: natural resources depletion \( (N_d) \) and Combustible renewables and waste \( (E_r) \). The particulate emission damage are calculated as productivity losses in the workforce due to premature death and illness. The variable thus represents society’s cost due to particulate emission damage that would be deducted from the adjusted net savings of a country which measures the value of a specified set of assets, excluding capital gains. Positive adjusted net saving indicate increasing social welfare (World Bank Development Data Group, 2010, 2014). Natural resource depletion is the sum of net forest depletion, energy depletion, and mineral depletion (ibid.). The third environment related variable measures to what extent biomass is used for energy production. Biomass is currently the most important renewable energy source, though it’s environmental benefits are often disputed (Field et al., 2008).

Finally, for changes in the latent variable, \( L \), the algorithm suggests a set of economic, demographic, political and environmental predictors. Among economic predictor the variable Net foreign assets \( (D) \) is suggested to be a very powerful predictor. It is the ”sum of foreign assets held by monetary authorities and deposit money banks, less their foreign liabilities” (World Bank Development Data Group, 2010, 2014), as such it measures the indebtedness of a country. Interestingly, this variable did not appear among strongest predictors for any of the three components of \( L \). The other economic predictors are GDP per capita \( (G) \) and gross national expenditure \( (G_e) \), which is the sum of household final consumption expenditure, general government final consumption expenditure and gross capital formation (World Bank Development Data Group, 2010, 2014).

The demographic variables include the age dependency ratio (% of working-age population) \( (A_d) \), which is the ”ratio of dependents – people younger 15 or older than 64 – to the working-age population – those ages 15-64” (World Bank Development Data Group, 2010, 2014) and the fertility rate. Fertility rate is generally a good proxy for the development stage of a country, while the age dependency ratio is potentially a proxy for the potential economic productivity of a population, as a low rate would indicate a majority of the population in the working age.

The three political variables that the algorithm picked as good predictors are independence of judiciary system \( (J) \), women’s economic rights \( (R_f) \) and religious freedom \( (R_f) \). None of these three indicators appeared as predictors for the three components, Child Mortality, Education and CO2, though they are correlated with some of the political variables that were chosen as predictors for the components. All three variables are provided by the CIRI Human Right Data Project. The independence of judiciary system measures the extent to which the judiciary is independent of control from government bodies and the military etc. (Cingranelli et al., 2013). Women’s economic rights indicate to what extent a set of rights, such as equal pay for equal work, free choice of profession or employment and the right to gainful employment without the need to obtain a husband or male relative’s consent, equality in hiring and promotion practices, job security (maternity leave, etc.), non-discrimination by employers, the right to be free from sexual harassment in the workplace, the right to work at night, the right to work in occupations classified as dangerous and the right to work in the military and police force, are implemented (ibid.). Freedom of religion is finally a variables that indicates to what extent the freedom of citizens to exercise and practice their religious beliefs and convert to other religious beliefs is restricted by the government (ibid.).

The two environmental predictors that were selected by the algorithm are Adjusted savings: natural resources depletion \( (N_d) \), the sum of net forest depletion, energy depletion and mineral depletion and Adjusted savings: net forest depletion \( (F_s) \), giving forest depletion a particular important role as an environmental indicator. The two other variables that have been identified as relevant predictors as well, were mobile phone subscriptions \( (M_p) \) per 1000 inhabitants, indicating modern technology progress in a society and health expenditure per capita (% of GDP) \( (H) \).

A3.3 Dynamical Systems Modelling Results
Table A2: Dynamical System Models for the four dependent variables $CM, Ed, CO2, L$. The second set of $L$ models refers to models from the Bayesian model combination procedure

| Dependent Variable | Models | Bayes Factor |
|--------------------|--------|--------------|
| Child Mortality    | $\frac{\Delta CM}{CM} = -0.51 \frac{\Delta t}{t}$ | -398.23 |
|                    | $\frac{\Delta CM}{CM} = 0.01 \frac{\Delta t}{t} - 0.62 G$ | -374.63 |
|                    | $\frac{\Delta CM}{CM} = 0.19 M - 0.002 F^2 - 1.4 \frac{\Delta t}{t}$ | -363.62 |
|                    | $\frac{\Delta CM}{CM} = -0.03 T/G + 0.86 M - 6.4 \frac{\Delta t}{t} - 0.001 F^3$ | -363.56 |
|                    | $\frac{\Delta CM}{CM} = 0.19 M - 0.002 F^2 - 1.4 \frac{\Delta t}{t} + 0.44 \frac{\Delta t}{t} + 0.91 MF_t$ | -428.85 |
|                    | $\frac{\Delta CM}{CM} = 0.31 M - 0.002 F^2 - 1.67 \frac{\Delta t}{t} - 0.01 T/G + 0.42 \frac{\Delta t}{t} - 0.87 MF_t$ | -438.11 |
| Education          | $\frac{\Delta Ed}{Ed} = -0.019 W_t$ | -756.13 |
|                    | $\frac{\Delta Ed}{Ed} = 0.0002 CG - 0.1 W_t$ | -754.14 |
|                    | $\frac{\Delta Ed}{Ed} = 0.0003 CG - 0.003G - 0.01 W_t$ | -749.55 |
|                    | $\frac{\Delta Ed}{Ed} = -0.01 G + 0.01 CG - 0.03 W_t + 0.16 \frac{\Delta t}{t}$ | -747.98 |
|                    | $\frac{\Delta Ed}{Ed} = 0.001 CG - 0.01 G - 0.04 W_t + 0.25 \frac{\Delta t}{t} - 0.001GF_t$ | -750.11 |
|                    | $\frac{\Delta Ed}{Ed} = 0.001 G + 0.01 CG - 0.01 G - 0.04 W_t + 0.25 \frac{\Delta t}{t} - 0.001GF_t$ | -750.59 |
| CO2 Emissions      | $\frac{\Delta CO2}{CO2} = 4e^{-0.05} \frac{\Delta t}{t}$ | -437.50 |
|                    | $\frac{\Delta CO2}{CO2} = 2.6e^{-0.05} \frac{\Delta t}{t} + 0.003 N_t E_m$ | -409.30 |
|                    | $\frac{\Delta CO2}{CO2} = 2.4e^{-0.05} \frac{\Delta t}{t} + 0.004 N_t E_m - 0.004 \frac{\Delta t}{t}$ | -399.20 |
|                    | $\frac{\Delta CO2}{CO2} = 2.2e^{-0.05} \frac{\Delta t}{t} - 4.3e^{-0.05} G^2 + 0.005 N_t E_m - 0.003 \frac{\Delta t}{t}$ | -401.57 |
|                    | $\frac{\Delta CO2}{CO2} = 1.9e^{-0.05} \frac{\Delta t}{t} + 0.07 G E_m - 0.08 C E_m + 0.002 D + 0.004 N_t C$ | -403.02 |
|                    | $\frac{\Delta CO2}{CO2} = 2.3e^{-0.05} \frac{\Delta t}{t} - 0.0004 G^3 + 0.11 G E_m - 0.11 C E_m + 0.004 G C - 0.003 \frac{\Delta t}{t}$ | -393.57 |
| Latent Variable    | $\frac{\Delta \Psi}{\Psi} = -0.47 \frac{\Delta t}{t}$ | -688.78 |
|                    | $\frac{\Delta \Psi}{\Psi} = -0.41 \frac{\Delta t}{t} - 0.02 \frac{\Delta t}{t}$ | -678.68 |
|                    | $\frac{\Delta \Psi}{\Psi} = -0.02 G^2 + 0.02 G^2 - 0.002 \frac{\Delta t}{t}$ | -671.58 |
|                    | $\frac{\Delta \Psi}{\Psi} = 0.002 G^3 - 0.06 D F_t + 0.02 F^2 + 0.002 \frac{\Delta t}{t}$ | -659.82 |
|                    | $\frac{\Delta \Psi}{\Psi} = 4.4 \frac{\Delta t}{t} - 1.2 D + 0.08 D F_t - 0.01 D F_t + 0.01 \frac{\Delta t}{t}$ | -652.53 |
|                    | $\frac{\Delta \Psi}{\Psi} = 0.46 \frac{\Delta t}{t} + 0.02 G^3 - 0.02 G^2 - 0.01 D F_t - 0.06 \frac{\Delta t}{t} - 0.02 N_t$ | -637.88 |
| Latent Variable    | $\frac{\Delta \Psi}{\Psi} = -0.33 W_t^2$ | -880.07 |
|                    | $\frac{\Delta \Psi}{\Psi} = -0.33 W_t^2 - 9.7e^{-0.6} \frac{\Delta t}{t}$ | -848.98 |
|                    | $\frac{\Delta \Psi}{\Psi} = -0.05 G + 0.01 C G - 0.2 W_t^2$ | -856.86 |
|                    | $\frac{\Delta \Psi}{\Psi} = -0.08 G + 0.01 C G - 0.4 W_t^2 + 1.7 \frac{\Delta t}{t}$ | -841.49 |
|                    | $\frac{\Delta \Psi}{\Psi} = -0.07 G + 0.01 C G - 0.4 W_t^2 + 1.7 \frac{\Delta t}{t} + 0.8e^{-0.6} \frac{\Delta t}{t}$ | -820.78 |
|                    | $\frac{\Delta \Psi}{\Psi} = 0.01 G C - 0.07 G + 2.1 \frac{\Delta t}{t} - 0.46 W_t^2 - 2e^{-0.4} F^3 + 7.1e^{-0.9} \frac{\Delta t}{t}$ | -816.06 |
