Economic Modelling 107 (2022) 105717

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

National development delivers: And how! And how?

Lant Pritchett

Oxford Blavatnik School of Government and Center for International Development, Harvard Kennedy School, USA

ARTICLE INFO

Keywords:
Economic development
Economic growth
State capability
Human wellbeing

ABSTRACT

National development is empirically necessary and sufficient for high levels of human wellbeing. Measures of three elements of national development: productive economy, capable administration, and responsive state, explain (essentially) *all* of the cross-national variation in the Social Progress Index (SPI), an omnibus indicator built from 58 non-economic indicators of human wellbeing. *How* national development delivers on human wellbeing varies, in three ways. One, economic growth is much more important for achieving wellbeing at low versus high levels of income. Two, economic growth matters more for "basic needs" than for other dimensions of wellbeing (like social inclusiveness or environmental quality). Three, state capability matters more for wellbeing outcomes dependent on public production. These findings highlight the key role of national development—and particularly economic growth—as instrumental to increased human wellbeing, which is increasingly challenged in favor of "small" programmatic and project design which is, at best, of third order of importance.

There are widespread critiques that "mainstream" development placed too much priority on economic growth. Many multi-lateral and bilateral development agencies and philanthropic and civil society actors have shifted from promoting national development as four-fold transformation towards small-bore, programmatic and project efforts to mitigate the consequences of the lack of development on specific indicators or on targeted groups (Banerjee and Duflo, 2011; Pritchett, 2015). For instance, the mission statement of the development advocacy group, Social Progress Imperative is[1]:

> We dream of a world in which people come first. A world where families are safe, healthy and free. Economic development is important, but strong economies alone do not guarantee strong societies. If people lack the most basic human necessities, the building blocks to improve their quality of life, a healthy environment and the opportunity to reach their full potential, a society is failing no matter what the economic numbers say. The Social Progress Index is a new way to define the success of our societies. It is a comprehensive measure of real quality of life, independent of economic indicators.

Placing less emphasis on national development on the premise that it does not reliably deliver on improvements in human wellbeing is empirically wrong. Suppose one ignores the "economic numbers" entirely[2] and takes the Social Progress Index (SPI) and its three major components–Basic Human Needs, Foundations of Wellbeing, and Opportunity–as the normative goals for improved human wellbeing. National development, measured by empirical proxies for three of its components; GDP per capita, State Capability, and Democracy,[3] is a strongly necessary and sufficient condition for achieving high levels for each of these four non-economic measures of wellbeing. There are *no* countries with high levels of *any* of the four SPI goals with low national development (empirical necessity). There are *no* countries with high levels of national development with low levels on *any* of the four SPI indicators (empirical sufficiency).

An examination of the connections between the three empirical measures of national development (GDPPC, State Capability, and Democracy) and the detailed physical indicators of wellbeing used in the SPI and its three major components and produces three additional

E-mail address: lant_pritchett@harvard.edu.

[1] I would like to thank Andres Velasco for conversations that pushed this gestating idea into birth. I would also like to thank Addison Lewis for excellent research support and the participants at a seminar hosted by Economic Modelling and the anonymous referee for helpful comments.

[2] Among the "economic numbers" the Social Progress Index ignores are all income/consumption measures of poverty. But it is widely accepted that cross-national variation in the level and change in poverty (at all the commonly used poverty lines) is almost entirely accounted for by the level and change in median (or mean) income/consumption which is tightly correlated with having a productive economy (Pritchett, 2020; Dollar et al., 2015; Dollar et al., 2016; McKenzie, 2020).

[3] I do not include a measure of the element of national development reflecting a shared identity and equality of treatment as I have yet to find a suitably reliable and general measure of the concept that does not lead to circularity in explaining wellbeing outcomes—e.g. using measures of inequality of income/consumption to explain outcomes that depend on inequality–not because I think this dimension of national development is less important.

https://doi.org/10.1016/j.econmod.2021.105717
Received 30 June 2021; Received in revised form 26 November 2021; Accepted 30 November 2021
Available online 7 December 2021

findings.

First, the impact of economic growth on aggregate human wellbeing is tremendously more important for poorer than richer countries. Across a range of flexible specifications (quartic, splines at quintiles) the elasticity of SPI with respect to GDPPC is highest at GDPPC between P$5000–10,000 (near the level of India, Vietnam, Nigeria) and then falls to (roughly) zero by a level of GDPPC of P$25,000–30,000 (near the level of Turkey or Greece).

Second, GDPPC is generally much more important for elements of human wellbeing the SPI regards as "Basic Needs" (like nutrition and basic health, access to water and sanitation, improved shelter) than for more advanced elements in "Opportunity."

Third, if one separates the components of national development into "economic" (GDPPC) and "governance" (state capability and democracy) the data suggests that growth has a larger impact on elements of human wellbeing that are primarily "private" goods (like nutrition and shelter) whereas "governance" is more important for "public" goods—like public safety or the environment.

The concluding section suggests that the debate in development about the priority to national development (including economic growth) is largely a confusion from the richer countries not acknowledging the difference between their priorities and those of people, communities, and countries with GDPPC an order of magnitude lower than the level the citizens of rich countries enjoy.

1. National development and human wellbeing: concepts and data

My overall question is: "How strongly are country measures of *national development* related to aggregates of direct physical measures[4] of *human wellbeing* of people living in that country?" In addition, there are questions like: "How do these relationships vary across levels of national development?" and "Which elements of national development are related most strongly to which dimensions of human well-being?" This section describes the measures of three elements of national development (productive economy (section I. a), state capability (section I. b) and responsive government (section I. c)) and the measures of social progress (section I. d) that I use to address these questions.[5]

1.1. Productive economy: PPP GDP per capita

The productivity of a national economy is (crudely) captured by estimates of Gross Domestic Product per person. I use the Penn World Tables 10.0 data (Feenstra et al., 2015) which provide estimates of GDP in purchasing power exchange rates (for which I use "P$") where, in principle, a "dollar" of GDP represents similar purchasing power across countries. Despite assertions otherwise (even sometimes by people who should know better) professional economists have never confused GDP per capita, a measure of economic product and productivity, with a measure of human wellbeing.[6]

1.2. Capable administration: state capability

The modern state takes on a wide range of functions that create the legal, regulatory and policy realities and determine the possible scope of action by people and organizations within a country. This includes the *imposition of obligations* (e.g. collection of taxes, definition and prevention of crime, enforcement of contracts, regulation (health, environmental, safety, land-use, economic)) and the *provision of services* (e.g. health care, schooling, utilities, infrastructure) and *redistribution* (e.g. old age pensions, safety nets, insurance). By "state capability" I mean the extent to which the organizations tasked with these tasks are able in practice to carry out and implement their duties and functions in ways that promote the organization's stated objectives. Overall state capability is an aggregation and capabilities of specific organizations may vary strongly within a country (Andrews et al., 2016).

My empirical indicator of state capability is the simple average of four indicators reported by the World Governance Indicators (WGI) (Kaufmann, Kraay and Mastruzzi, 2005): *Rule of Law*, *Government Effectiveness, Regulatory Quality,* and *Control of Corruption.* These four are very highly correlated amongst themselves (the bivariate correlations are all above .9) but these four are conceptually and empirically distinct from the other two indicators in the WGI: *Voice and Accountability* and *Political Stability/Absence of Violence/Terrorism*.[7] The WGI SC is normed from 1 (worst rated country) to 100 (best rated) country.

1.3. Responsive government: POLITY(K) cumulated democracy/autocracy

I don't really want to conflate the conceptually broad historical political transformation to a "responsive government" with the much narrower notion of "democracy"–but I am going to. The POLITY ranking typically gives each country in each year a score on autocracy from 0 to 10 and on democracy from 0 to 10 and the Democracy score minus the Autocracy score provides a score that ranges from −10 to 10. In addition, the POLITY provides special codes when countries are dominated by a foreign power or in a civil war or conflict of −66, −77, or −88.

Two non-standard transformations create a variable that represents a "stock" rather than a current "flow" of democracy, which I call POLITY(K). First, I transform all years of −66, −77, and −88 into a −10, assuming these years of instability/conflict contribute to a country's "stock" of democracy the same as a year of complete autocracy (−10). Second, I take a time-weighted average of all available POLITY2 scores using a discount factor of $\delta = 0.05$ so that weight of the contribution to the cumulated "stock" of POLITY in year T of POLITY observed in year T-t is $w_t=(1-\delta)^{(T-t)}/\text{sum}(w_t)$. I rescale POLITY(K) to a minimum of 1 (worst country) and maximum of 100 (best country).

I prefer the stock measure for two reasons. One, I think the time-weighted average better represents a measure the "institutions" of democracy as the longer they last the more "heft" in determining outcomes democratic practices have.[8] Two, in assessing the impact of a "responsive government" on wellbeing outcomes the effects are likely to have very long lags and there is little reason to expect that, say, the proportion of

[4] By relying on *physical* measures (e.g. access to water, homicide rate, schooling, health) I am bracketing the question of the relationship between physical conditions (e.g. access to sanitation) and "psychic" or "subjective" wellbeing like self-reported measures of happiness or life satisfaction.

[5] Measures of *national* development are not even on the same *ontological* basis as national measures of human wellbeing. For the latter the individual (or household) is the *ontological* unit and aggregation is secondary. As a simple example, one can measure each person's height. One could then aggregate individual heights into the average height of people in Nepal and the average height of people in Kenya, but one could also talk equally well of all kinds of aggregations: average height of left-handed versus right-handed, average height of people born on a Tuesday versus people born on Wednesday, average height of people whose names start with A-K versus L-Z. Height is ontologically individualized and aggregation is secondary. In contrast, countries have characteristics that are not the simple aggregation of ontologically individualizable characteristics of its citizens/residents.

[6] In 1934 Simon Kuznets, the pioneer of national accounting and GDP measurement, said: "The welfare of a nation can scarcely be inferred from a measure of national income."

[7] Factor analysis of the huge array of "governance" indicators suggests that "state capability" can be distinguished from other elements of governance–like democracy or human rights or political stability (Drumm, 2015).

[8] For instance, a review of the literature on political instability between 1955 and 2002 by a task force of a dozen scholars (Goldstone and Ulfedler, 2004) found that "the key to maintaining stability appears to lie in the development of democratic institutions that promote fair and open competition, avoid political polarization and factionalism, and impose substantial constraints on executive authority."

people with sanitation or gender parity in secondary education are going to respond fully and instantaneously to an election or a coup. Pakistan provides one example of the volatility of POLITY2 and the implications of this smoothing. Pakistan's POLITY2 scores falls from 8 to −7 from 1976 to 1977 after Zia-ul-Haq declared martial law and became president. It then rose from −4 in 1987 to 8 in 1988 with the election of Benazir Bhutto. It then fell again to −6 in 1999 when the military again took power then rose from −5 in 2006 to 5 in 2007 and further to 7 by 2018. If we took just the conditions in a given year then in 2018 (the latest data) Pakistan's score would be a 7, India's a 9 and Denmark's a 10. In the POLITY(K) measure in 2018 Pakistan is 52, India 85, and Denmark 100, as the stock measure reflects Pakistan's more volatile political past than India.

Before the reader gets caught up in articulating the very many ways in which these three empirical measures are weak counter-parts of their conceptual counter-parts, keep in mind the basic empirical finding below is that these three indicators explain essentially *all* of the variation across countries in SPI. As measurement error generally attenuates results and weakens goodness of fit, better indicators would likely make the already very strong findings even stronger. By taking the simple and easily available indicators roughly "as is" I err on the side of not being susceptible to suspicions of "data mining."

1.4. The relationships among the components of national development

Fig. 1 (a, b and c) show the relationships amongst productive economy (GDPPC), state capability (WGI SC), and responsiveness of the state (POLITY(K)), as their correlations are essential to understanding the empirical results below.

GDPPC and WGI SC are strongly correlated, $\rho = .794$, though with notable exceptions (Fig. 1a). Countries dependent on "point source" natural resources tend to have low levels of state capability for their income.[9] Along the 'southeast' edge of the relationship (low capability for a given GDPPC) one sees United Arab Emirates (ARE), Bahrain (BHR), Saudi Arabia (SAU), Kuwait (KWT), Russia (RUS), Gabon (GAB), Iraq (IRQ), and two countries with very low WGI SC even at high GDPPC: Equatorial Guinea (GNQ) and Turkmenistan (TKM) (in the graph these two countries overlap). Along the "northwest" of the relationship there are countries with exceptional strong WGI SC for their income: Rwanda (RWA), Georgia (GEO), and Chile (CHL).

POLITY(K) is much less highly correlated with GDPPC, $\rho = .494$, with three distinct patterns (Fig. 1b). One, as is well attested in the literature, at very low levels of income there are few stable democracies as many poorer countries have had periods of democratic and non-democratic politics (Goldstone and Ulfelder, 2004). Two, in the middle range of GDPPC there is an extremely wide variation as there are countries at similar levels of GDPPC at high levels (e.g. India (IND) 85.4 and Jamaica (JAM) 99), middle levels (e.g. the Philippines (PHL) at 56.7) and low levels (e.g. Morocco (MAR) 12.5) of POLITY(K). Three, at the high levels of income there are two clear groups, with 17 high income countries at or near POLITY(K) of 100–in the graph these country names are one blur along the top edge—and the high-income oil countries (e.g., Kuwait (KWT), United Arab Emirates (ARE)) and Singapore (SGP) with high GDPPC and low POLITY(K). There are no countries very high GDPPC and middling POLITY(K).

WGI SC and POLITY(K) are moderately correlated, $\rho = .669$ (Fig. 1c), with a similar pattern to GDPPC and POLITY(K) that in the middle range of state capability one sees the entire range of possible ratings of POLITY(K), and vice versa. At the middle range of democracy there are countries with both very low and very high WGI SC: Singapore has POLITY(K) of 33 and WGI SC of 100 whereas Sudan (SDN) has an almost identical POLITY(K) at 31 and WGI SC of only 4.4.

This array of correlations illustrated in Fig. 1a,b, and 1c makes multivariate regressions both necessary and possible. The correlations amongst these three variables are so high that bivariate correlations would be highly misleading: a bivariate regression of human wellbeing on GDPPC alone would (implicitly) attribute the impacts of higher state capability on and democracy on wellbeing to GDPPC. Yet the correlations are not so high that multi-collinearity makes it in principle impossible to distinguish with precision amongst these three conceptually distinct elements of national development.

1.5. Human wellbeing: Social Progress Index and indicators

The Social Progress Index, its sub-indices, and its component indicators, have two major advantages in an examination of the empirical association between national development and measures of human wellbeing.

First, as seen in the introduction, their stated purpose is to displace traditional economic measures. The Social Progress Initiative has put forward a set of normative and evaluative measures as alternatives to economic performance. The SPI not only does use GDP per capita, it does not use *any* economic measure, even those based on directly measured household income or consumption: not headcount poverty measures, not average/median consumption, not consumption adjusted for inequality (e.g. normative Atkinson indexes).

Second, the SPI has a much broader array of indicators than have been previously examined. The UN Human Development Index (HDI) incorporated health and education measures since the 1990s. The connection between specific indicators of well-being (especially measures of country health outcomes like under 5 child mortality and life expectancy and education measures) and GDP per capita has been extensively studied[10]. The present paper extends Pritchett and Kenny (2013) which was limited to poverty, health (under 5 mortality), and education (years of schooling). Empirical explorations with other cross-national data with multiple indicators of wellbeing (such as the Legatum Prosperity Index) find very similar empirical results (Pritchett and Lewis, 2021).

Third, I take advantage of the Social Progress aggregation of many indicators into larger components. The cross-national correlates, including GDPPC, with many other elements of human wellbeing, such as the extent of schooling, the frequency of various forms of malnutrition, access to various types of infrastructure (sanitation, electricity, transport), have also been studied extensively, Easterly (1999), for instance, examines the correlation of the level and changes in GDP for 81 distinct indicators. Using a single data source and proceeding from aggregate to sub-aggregate to components using the exact same regression specification has the advantage of comparability across indicators.

The Social Progress Index (SPI) (Table 1) is an average of three components, which are each the average of four sub-components, and the sub-components are constructed from individual elements.

The component ***Basic Human Needs*** (***B***) is the average of the four sub-components (i) *Nutrition and Basic Medical Care (B:NB), (ii) Water and*

[9] This distinguishes between countries/economies dependent on "diffuse" natural resources (like stable crop agriculture or livestock) versus those dependent on "point source" natural resources that are not geographically dispersed (like oil, gold or diamonds) on the conjecture that these two types of natural resources create very different political dynamics as the potential for extracting rents is entirely different (Isham, Pritchett, Woolcock and Busby, 2005).

[10] The "Preston curve" (1975) relationship between health and GDPPC has inspired a massive literature to which I have occasionally contributed in ways that both inspire and amplify the points in this paper. Filmer and Pritchett (1999) show that essentially all of the variation in child mortality are explained by a few country-level factors, far and away the most important being GDPPC). This other work also addresses technical issues this paper does not explore, like causality (Pritchett and Summers, 1996) and the dynamics of the relationship (Pritchett and Viarengo, 2010).

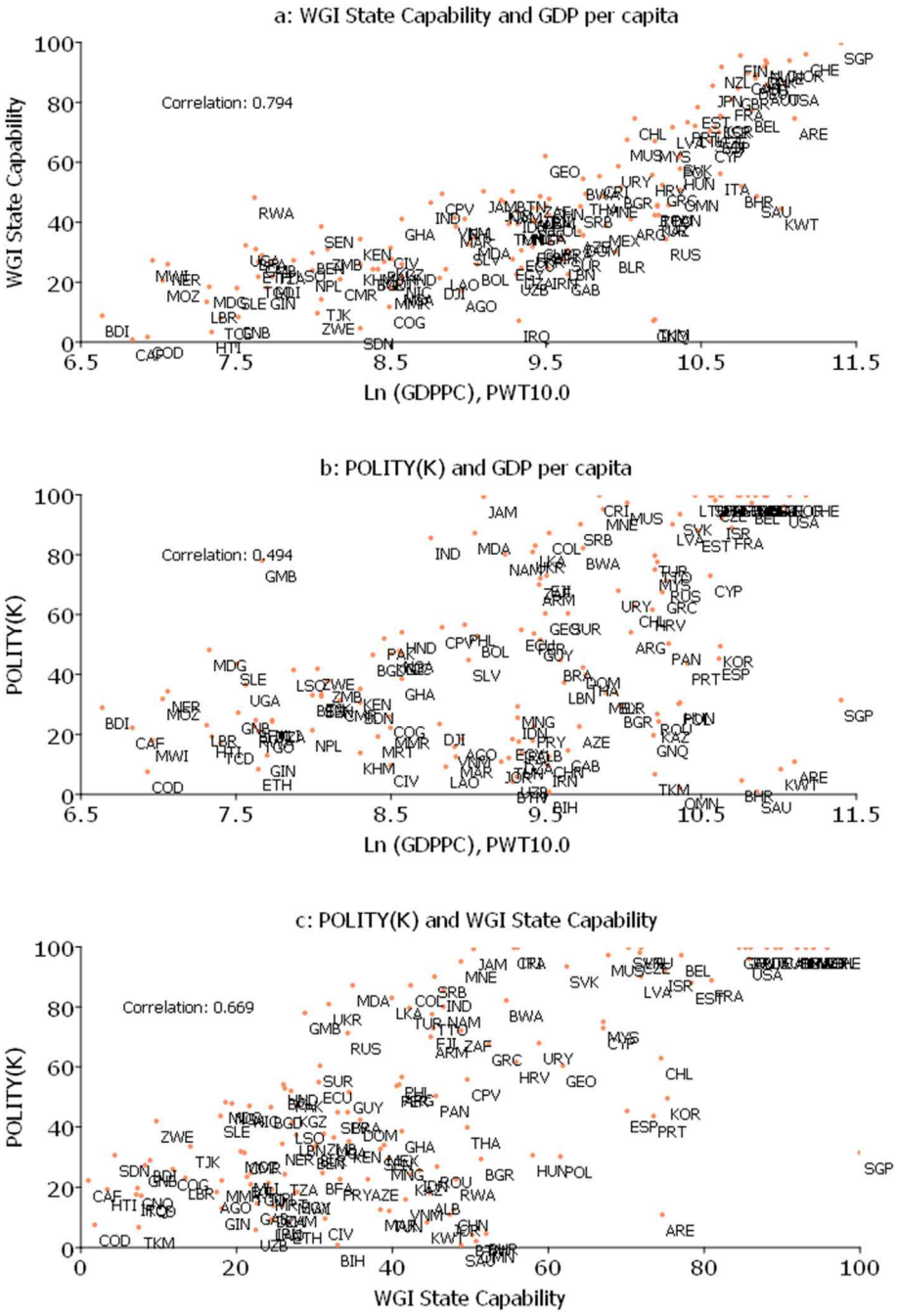

Fig. 1. a: WGI State Capability and GDP per capita b: POLITY(K) and GDP per capita c: POLITY(K) and WGI State Capability.

Sanitation (B:WS), (iii) Shelter (B:HS), and (iv) *Personal Safety (B:SF).*

The SPI component ***Foundations of Wellbeing*** (F) is the average of the sub-components (i) *Access to Basic Knowledge (F:BK), (ii) Access to Information and Communications (F:IC), (ii) Health and Wellness (F:HW),* and (iv) *Environmental Quality (F:EQ).*

The SPI component ***Opportunity*** **(O)** is the average of the sub-components: (i) P*ersonal Rights (O:RP), (ii) Personal Freedom of Choice (O:FP), (iii) Inclusiveness (O:IV),* and (iv) *Access of Advanced Education (O:AE).*

Each of the 12 sub-components is built from a collection of 50 objective and subjective indicators such as: percent of population undernourished, homicide rate (deaths per 100,000), and perceived criminality on a 1 to 5 scale.

To make the regression results comparable I re-norm all of the SPI variables (components, sub-components, indicators) to a 1 (lowest country on that particular index) to 100 (highest country on that particular index) scale.

Any aggregate index built from multiple indicators in physical units

Table 1
The Social Progress Index, its three components, their four sub-components, and the elements of the sub-components.

1	3	12 sub-components	58 indicator elements of each sub-component
Social Progress Index	Basic Human Needs	Nutrition and Basic Medical Care (NB)	Undernourishment (% of pop.), Deaths from infectious diseases (deaths/100,000), Child stunting (% of children), Maternal mortality rate (deaths/100,000 live births), Child mortality rate (deaths/1000 live births)
		Water and Sanitation (WS)	Unsafe water, sanitation and hygiene attributable deaths (per 100,000 pop.), Populations using unsafe or unimproved water sources (%), Populations using unsafe or unimproved sanitation (%)
		Shelter (HS)	Usage of clean fuels and technology for cooking (% of pop.), Access to electricity (% of pop.), Household air pollution attributable deaths (deaths/100,000)
		Personal Safety (SF)	Traffic deaths (deaths/100,000), Political killings and torture (0 = low freedom; 1 = high freedom), Perceived criminality (1 = low; 5 = high), Homicide rate (deaths/100,000)
	Foundations of Wellbeing	Access to Basic Knowledge (BK)	Access to quality education (0 = unequal; 4 = equal), Women with no schooling, Gender parity in secondary attainment (distance from parity), Primary school enrollment (% of children), Secondary school attainment (% of population)
		Access to Information and Communications (IC)	Access to online governance (0 = low; 1 = high), Media censorship (0 = frequent; 4 = rare), Internet users (% of pop.), Mobile telephone subscriptions (subscriptions/100 people)
		Health and Wellness (HW)	Access to quality healthcare (0 = unequal; 4 = equal), Access to essential services (0 = none; 100 = full coverage), Premature deaths from non-communicable diseases (deaths/100,000), Life expectancy at 60 (years)
		Environmental Quality (EQ)	Greenhouse gas emissions (total CO2 equivalents), Particulate matter, Biome protection, Outdoor air pollution attributable deaths (deaths/100,000)
	Opportunity	Personal Rights (RP)	Political rights (0 = no rights; 40 = full rights), Freedom of expression (0 = no freedom; 1 = full freedom), Freedom of religion (0 = no freedom; 4 = full freedom), Access to justice (0 = non-existent; 1 = observed)
		Personal Freedom and Choice (FP)	Property rights for women (0 = no right; 5 = full rights), Vulnerable employment (% of employees), Corruption (0 = high; 100 = low), Early marriage (% of women), Satisfied demand for contraception (% of women)
		Inclusiveness (IV)	Equality of political power by socioeconomic position (0 = unequal power; 4 = equal power), Equality of political power by social group (0 = unequal power; 4 = equal power), Equality of political power by gender (0 = unequal power; 4 = equal power), Discrimination and violence against minorities (0 = low; 10 = high), Acceptance of gays and lesbians (0 = low; 100 = high)
		Access to Advanced Education (AE)	Quality weighted universities (points), Citable documents, Women with advanced education (%), Years of tertiary schooling

Source: https://www.socialprogress.org/.

suffers from the impossible challenge of choosing weights.[11] There is no consensus resolution to questions like: "a one unit decrease in 'outdoor air pollution' produces the same increase in a country's overall wellbeing as an X unit increase in 'access to electricity', a Y unit reduction in 'maternal mortality' or a Z unit increase in 'property rights for women.'?" as the weights X, Y, and Z are a subjective value judgments. This intractability typically leads to a "focal point" solution like equal weights, but using equal weights is due to the *lack* of any justification of any set of weights, *not* a justification of equal weights.

However, for omnibus indicators, like the overall SPI, the weights question is not empirically that important due to the high inter-correlations amongst the many components, sub-components, and elements (which I return to in Section IV.C).[12] The SPI and its three components; Basic Human Needs, Fundamentals of Wellbeing and Opportunity are also quite highly correlated amongst themselves (see Graph A.1 in the Graphical Appendix). On the other hand, the correlations are, naturally, lower amongst the 12 sub-components and lower still among the 50 underlying indicators. This lack of perfect correlation allows different measures of wellbeing to have revealingly distinct patterns of association with national development and its components, as we will see in Section IV.

1.6. Estimation method and functional form

All estimates are OLS of the *level* of an indicator of wellbeing on the three indicators of national development for a single year, 2018. Below I discuss why I neither attempt to use any technique to resolve causality nor to use the time series variation.

All variables, both of national development and of wellbeing are each scaled from 1 to 100. I allow that the association/elasticity of wellbeing

[11] Any economic aggregate, like total consumption or GDP faces the same issue of (literally) adding up apples and oranges. Economics has a coherent theory-based rationale for why the aggregate "total value of fruit" adds apples and oranges using prices as weights, an approach with strengths and weaknesses. It generally works well—or at least is well understood–for ordinary private goods but economics has long acknowledged the difficulty of valuation of public goods/bads (non-rival and non-excludable) and externalities (which are public goods/bads bundled with private goods) which lack markets in which prices are determined. The many justified criticisms of GDP and the many "corrections" to it for, say, the environment, via correcting the undercounting of the depletion of goods (for genuine savings rates (Clemens and Hamilton, 1999) or the production of bads (e.g. air pollution) or both. There are many efforts into green national accounts and reviews of those efforts, see, *inter alia* Narloch et al., (2016), Li and Lofgren, 2011, which adjust the distortions in "true" concepts from using market prices for GDP which ignore externalities and depletion of natural resource stocks.

[12] As an illustration, principal components is a data reduction technique commonly used for creating an index from multiple elements as this procedure creates weights for an ordered set of principal components that (intuitively, if not mathematically precisely) maximize the common variation among a set of variables. The correlation between the first principal component of the 12 sub-components and the SPI is .999 and the fraction of variance among the 12 sub-components explained by the first principal component is .96 and the deviation of the principal components derived weights from equal weights is not empirically that large.

Table 2
OLS regressions Social Progress Index and its three components on measures of national development.

Variable	Regression Statistic	SPI	Basic Human Needs	Fundamentals of Well-Being	Opportunity
WGI SC	Coefficient	0.547	0.360	0.544	0.694
	Std Error	0.061	0.065	0.077	0.078
	p-level	1.42E-15	1.30E-07	5.82E-11	3.58E-15
POLITY(K)	Coefficient	0.081	−0.022	0.092	0.187
	Std Error	0.029	0.031	0.037	0.038
	p-level	0.0062	0.4807	0.0136	0.0000
GDPPC (1–100 scale)	Coefficient	3.227	5.247	2.129	1.188
	Std Error	0.376	0.401	0.474	0.485
	p-level	1.69E-14	6.81E-26	1.51E-05	1.56E-02
GDPPC Squared	Coefficient	−0.102	−0.176	−0.054	−0.037
	Std Error	0.018	0.020	0.023	0.024
	p-level	1.46E-07	1.79E-15	2.25E-02	1.23E-01
GDPPC Cubed	Coefficient	1.26E-03	2.32E-03	5.25E-04	4.28E-04
	Std Error	3.18E-04	3.40E-04	4.02E-04	4.11E-04
	p-level	0.0001	0.0000	0.1942	0.2996
GDPPC Quartic	Coefficient	−5.51E-06	−1.04E-05	−1.87E-06	−1.88E-06
	Std Error	1.73E-06	1.85E-06	2.19E-06	2.24E-06
	p-level	0.0019	0.0000	0.3957	0.4036
Number of countries		145	145	145	145
R-Squared		0.894	0.874	0.846	0.859
Std. Dev. Residual		7.9	8.5	10.0	10.3
F test GDPPC linear		4.93E-21	4.84E-32	2.41E-10	4.93E-04
F test WGI SC linear		0.1214	0.0824	0.4243	0.3565
F test PK linear		0.0024	0.2641	0.0082	0.0001

indicators with respect to GDPPC to be non-linear (and the partial scatterplots in GA.2 show strong non-linearity) using either a quartic in GDPPC (in the results in Table 2) or a spline. The non-linearity at the very upper end of GDPPC creates difficulties in the higher order polynomial terms from outliers at the upper end of GDPPC (which is irrelevant to the range of incomes I am interested in) and hence I excluded from all estimates three small, very high GDPPC countries: Luxemburg, Ireland, and Qatar. This results in 145 countries used in all regressions.

2. National development delivers: and how!

The first empirical finding, reported in Table 2 and illustrated in Fig. 2 Fig. 3 is that the SPI and each of the three components are very strongly associated with all three components of national development. The

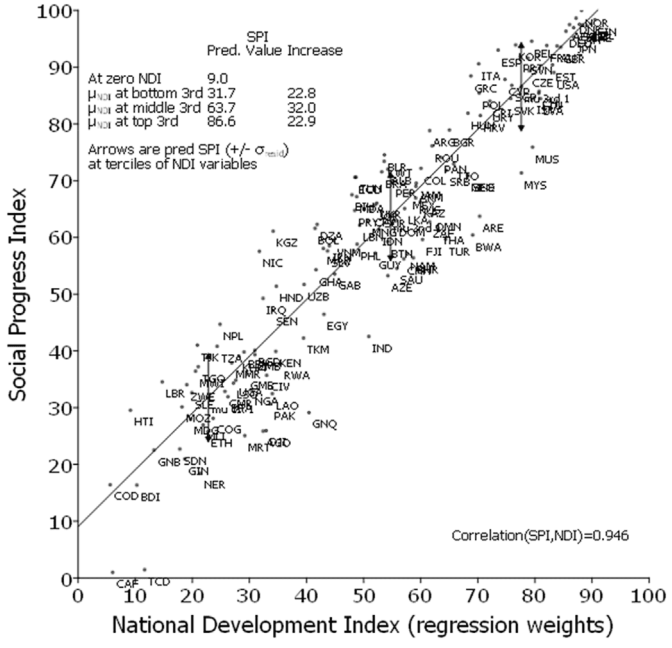

Fig. 2. Social Progress Index and National Development Index, at terciles. Source: Author's calculations.

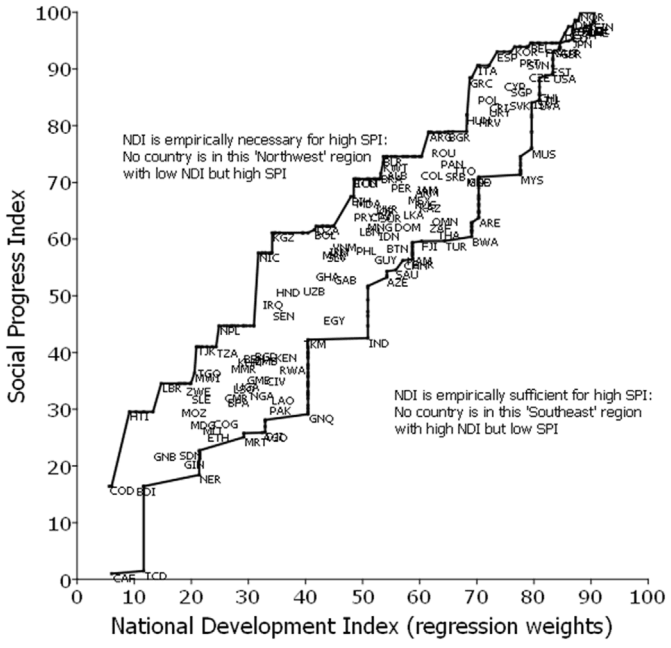

Fig. 3. National Development and SPI, necessary and sufficient.

graphs in the Graphical Appendix, GA.2, GA.3 and GA.4 show the partial scatter plots for each of the three national development variables for SPI and its three components.

The R Squared shows that nearly 90 (R2 = 0.894) percent of the variation across countries in the SPI is associated with GDPPC, WGI SC, and POLITY(K). This implies the correlation of SPI with a national development index (the national development regression predicted value of SPI) is 0.95, which is a strikingly high correlation in cross-national data. The R2 is 0.846 or above for each of the three SPI components.

There has been justifiable attention in medicine and the social sciences to the "replication crisis" that results from relying on low statistical power results and the use of standard hypothesis test significance levels, like p = .05 or p = .01 in spite of the risks of small samples, data-mining, multiple tests, publication bias, etc. (Ioannidis, 2005; Camerer et al.,

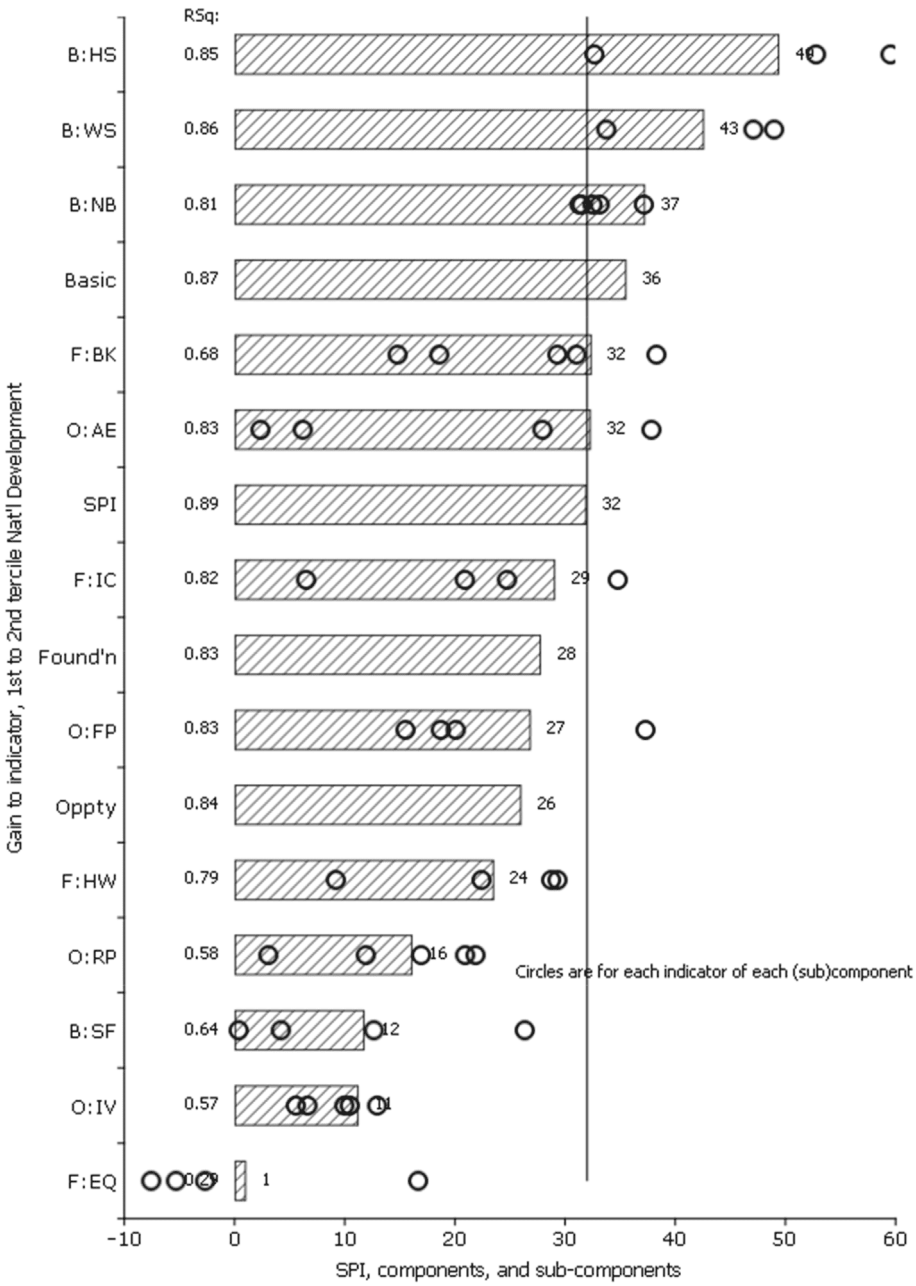

Fig. 4. Predicted increased in SPI, its three components, 12 sub-components and 50 indicators from an increase from the 1st to 2nd terciles on each element of national development. Source: Author's calculations.

2018; Young, 2019). The hypothesis test results for WGI SC and GDPPC (all terms) for SPI and each component are in the "six sigma" range or above and the p-levels are many orders of magnitude lower than the standard 1/5/10 percent levels. The results for POLITY(K) are much weaker, and while they do reject a zero coefficient (except for Basic Human Needs) they do so at more modest p-levels.

The F-tests of non-linear terms in GDPPC decisively reject linearity, again at p-levels orders of magnitude lower than 1 percent. There are mixed results for the linearity of WGI SC–never rejected at the 0.05 level–and POLITY(K) where there are less than 1 percent p-level rejections for three of the indicators.

Fig. 2 illustrates the association of SPI and national development (graphs for the other three SPI components are in Graphical Appendix, GA.5). The national development index for each country is the predicted value of SPI using the regression coefficients and the actual values of the three elements of national development: NDI (β_{SPI}^{OLS}) which, given the nature of OLS, is the linear index of national development which best predicts the SPI in mean square error. Increases in NDI are strongly and reliably associated with improvements in the SPI. The arrows in Fig. 2 show the predicted value of SPI plus/minus a residual standard deviation at each of the terciles of the NDI. At the average value of the bottom third of countries for each of WGI SC, POLITY(K) and GDPPC the NDI is 22.7 and the predicted SPI is 31.7. The standard deviation of the residual is 7.8 so most countries at an NDI of 22.7 would be expected to have a SPI between 23.9 (Ethiopia (ETH) is 25.8) and 39.5 (Tajikistan is 41.1).

Fig. 2 is the scatter plot of SPI and NDI. If a country moves from the mean of the bottom third to the mean of the second third of countries on each of NDI elements its NDI (predicted SPI) increases from 31.7 to 63.7—an increase of 32 points, nearly doubling. With the same residual standard deviation of 7.8 the expected range of SPI at that NDI would run from 55.9 Azerbaijan (AZE) is 53.2) to 71.5 (Kuwait (KWT) is 73.7). Even a country with low SPI outcomes at the middle of the second tercile on

NDI components would be expected to have much higher SPI than a country with high performance at a the bottom third NDI (55.9 versus 39.5).

Fig. 3 illustrates the tight association between national development (NDI) and human wellbeing (SPI) with a shape that includes all observations. The upper line in Fig. 3 is the upper envelope of SPI achievement for any country at a given level of NDI or lower. The lower line in Fig. 3 is the lowest SPI for any country with that level of NDI or lower. The space between the upper and lower lines is the "observed envelope" of SPI-NDI outcomes. The "white space" are combinations of SPI and NDI that do not occur among the 145 countries.

The blank portion in the upper-left ("northwest") of the graph shows that national development is *empirically necessary* for high overall levels of human wellbeing. There are no countries with high SPI with low NDI. Suppose a country aspired to the level of human wellbeing achieved by Argentina (ARG) of 78.8 (which is the high end of "developing" countries but lower than all (old) OECD countries). No country has ever achieved that level of SPI without getting to a level of national development of at least that of Argentina, 61.6.

Conversely, achieving high levels of national development is an empirically sufficient condition for achieving high SPI. India (IND) illustrates the lower line with an NDI of 50.9 but a SPI of 42.5 as every country with a higher NDI has a higher SPI. There is no country with a high NDI which does not achieve a high SPI.

I assert results illustrate national development is "strongly" empirically necessary and sufficient for SPI as the envelope is, visually and intuitively a small portion of the possible space. With an exact logical or mathematical "necessary and sufficient" the envelope would be a line (fill zero space) whereas if there were no association at all the envelope would roughly fill the space.

Fig. 4 shows the predicted impact of national development gains on the SPI, its 3 components, each of the 12 sub-components (bars) and the 50 constituent indicators within each sub-component (circles). Shown for each variable is the regression predicted increases from moving each of the three elements of national development (GDPPC, WGI SC, POLITY(K)) from the mean of the first tercile to the mean of the second tercile and, to the left of each indicator, the R2 of the national development regression.

Overall, the SPI increases by 32 points (indicated by the vertical line). Naturally, the gains are larger for some components and smaller for others. For three sub-components in the *Basic Needs* component: Shelter (B:HS), Water and Sanitation (B:WS), and Nutrition and Basic Health (B:NB) the gains are more than one for one from increases in national development. The predicted gain in *Basic Needs* from first to second tercile of NDI is 36. Two indicators related to education: Basic Knowledge (F:BK) and Access to Advanced Education (O:AE) also increase by 32 points.

There are four indicators with low predicted impact and low R2: Political Rights (O:RP), Personal Safety (B:SF)), Inclusiveness (O: IV), and Environmental Quality (F:EQ).

3. National development delivers: and how? Economic growth

In this section I illustrate two points: (i) that the elasticity of SPI with respect to GDPPC is highly non-linear and (ii) the pattern of responsiveness to GDPPC is different for different components of the SPI, with GDPPC more important for basics.

3.1. Non-linearity of impact of GDPPC on wellbeing

The regressions in Table 2 includes up to quartic terms in GDPPC to allow for the responsiveness of wellbeing (and of various components of wellbeing, which we explore more in Section IV) to vary across levels of income. As is the case with the share of food in total consumption (the Engel curve), one can easily imagine that at low levels of income increases in GDPPC are very important to wellbeing but, as the more pressing material needs are met, other components of well-being less responsive to GDPPC become more important. In the World Values

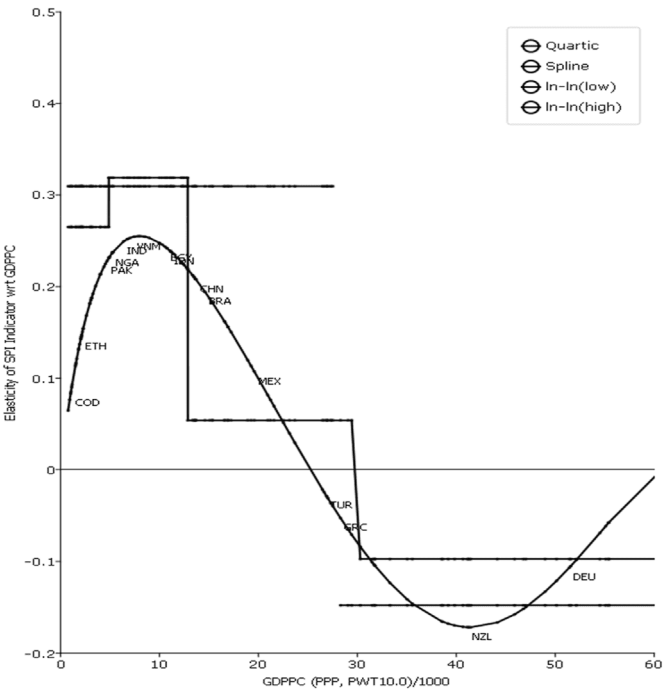

Fig. 5. The elasticity of SPI wrt to GDPPC, various functional forms. Source: Author's estimates.

surveys the proportion of people who are "post-materialist" (e.g. respond that economic growth is less of a priority) rises with the level of GDPPC (Inglehart, 2008; Pritchett, 2015, OWID).

This quartic functional form implies that the partial derivative of SPI wrt GDPPC varies as a cubic (equation 1) and hence in order to compare the "impact" across levels of GDPPC I calculate the elasticity of SPI wrt GDPPC (equation 2) for each level of GDPPC.

$$1)\ \frac{\partial SPI}{\partial GDPPC} = \beta_1 + 2\beta_2 GDPPC + 3\beta_3 * GDPPC^2 + 4\beta_4 * GDPPC^3$$

$$2)\ \varepsilon_{SPI,GDPPC} = \frac{\partial SPI}{\partial GDPPC} * \frac{GDPPC}{SPI}$$

Fig. 5 shows this quartic elasticity, which first rises as GDPPC increases, reaches a maximum of 0.255 at around P$7500 (roughly Morocco's level) and then falls, reaching zero at about P$25,000 (just above Chile) then goes further negative and then recovers. (The shape at very high levels of GDPPC is almost certainly an artefact of functional form: even a quartic polynomial can only be so squiggly.)

Fig. 5 also shows a functional form that allow splines, a continuous piecewise linear function but with "kinks" (discontinuous first derivatives). The splines are set at the upper range of the quartiles of GDPPC (P$4,868, P$12,850, P$29,420) and then the elasticity is calculated at the average GDPPC within each quartile (P$2,508, P$9,035, P$20,340, P$47,110). These "spline" elasticities in Fig. 5 show a very similar pattern to the quartic polynomial, with an elasticity at the lowest quartile of 0.26, rising to 0.32 for the second quartile, falling to 0.054 for the third and negative 0.097 for the highest quartile (the highest quartile, above P$ 29,420, are all "developed" countries). The population weighted average elasticity for countries with GDPPC less then P$28,000 with either the quartic or spline functional form is almost identical, 0.205 (quartic) and 0.207 (spline).[13]

[13] This changes if China is included or excluded only for the spline (as China is just above the threshold into the third quartile by GDPPC) whereas with the quartic the population weighted average is the same with or without China.

The simpler form for regression estimates of elasticities is to use double (natural) log functional form, and in Fig. 5 I estimate two completely separate regressions, divided at GDPPC of than P$28,000 (which puts Greece just above the threshold). As expected, the results for the two sets of countries are completely different, with an elasticity of SPI wrt GDPPC (controlling for WGI SC and POLITY(K)) of 0.31 for countries less than P$28,000 and a negative elasticity for those above that level.

3.2. Impact of GDPPC on the different SPI components

Fig. 6 shows the same pattern of impact across levels of GDPPC separately for each of the three components of the SPI: *Basic Human Needs*, *Foundations of Wellbeing*, and *Opportunity* using the spline at quartiles for each.

As income increases from very low levels of income, those in the bottom quartile of countries, ranging only up to P$4868 (the upper range is near Bangladesh and Cote d'Ivoire) the gains are primarily in *Basic Needs* (an elasticity of 0.45) and some in *Foundations* (0.14) and little or no impact (estimate is slightly negative, −0.03) on *Opportunity*.

In the next quartile of countries by GDPPC, from P$4868 to P$12,850 (which ranges from Myanmar and Nigeria and the lower end to Peru and Sri Lanka at the upper end) there are high and roughly equal elasticities for all three components (*Basic* 0.32, *Foundations* 0.29, *Opportunity* .26).

In the third quartile of countries by GDPPC (which ranges from South Africa and Iran at the lower end to Myanmar and Greece at the upper end) the elasticity wrt GDPPC are all lower, highest for *Foundations of Wellbeing* (0.11), less so for *Basic Needs* (0.05)–which have already at these levels of GDPPC have reached high levels–and again the income impact on *Opportunity* falls to about zero (−0.01).

This pattern of responsiveness to GDPPC is intuitive. In *Basic Needs* are indicators like percent of children malnourished, child mortality, households with sanitation, and access to electricity, whereas in the *Foundations of Wellbeing* are indicators like primary school enrollment, percent using the internet, and life expectancy at age 60. While all of these are clearly in household's preferences and we should expect that, as people's income increases, they will consume more of each of them, one can expect a household without an indoor toilet would prioritize getting a toilet and a household with a malnourished child or one at risk of dying might prioritize that over getting internet. The goal of national development is to create conditions so that fewer and fewer households have to make either of those type of hard choices. A stronger response of *Basic Needs* than *Foundations of Wellbeing* to expansions in GDPPC from very low levels of income, followed by a large responsive of both *Basics* and *Foundations* at the next level (second quartile), and then *Foundations* more than *Basics* (especially keeping in mind these are *elasticities* and hence are percentage changes, not absolute changes), makes intuitive sense.

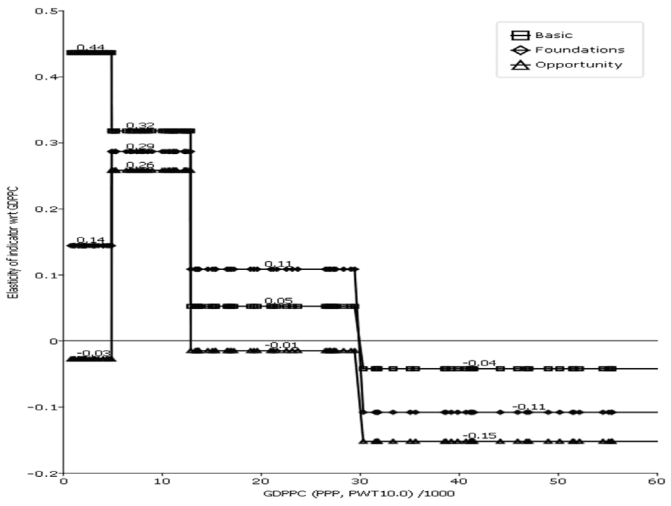

Fig. 6. Elasticity of three components of SPI wrt GDPPC, spline. Source: Author's calculations.

4. National development delivers: and how? Components of national development

While human wellbeing, especially in basics, rises strongly with GDPPC this is not suggesting that growth alone will solve all problems or that "development" means "growth." National development has always been more understood to be more than just income. This section examines how each of the three indicators of national development is associated with improvements in SPI and its components.

Given that the impacts of growth on these indicators is highly non-linear in order to compare the relative impacts of the elements of national development I compute the gains to SPI or its components from an increase across the terciles of the elements of national development. Fig. 7 shows the expected gains from moving each of the three national development indicators from the minimum possible value (normed to 1) to the average of the lst tercile, then the gain from the increase from the mean of the 1st tercile to the mean of the 2nd tercile, and from mean of the 2nd tercile to the mean of the 3rd tercile. For state capability the mean of the 1st tercile is 19, roughly the level of Madagascar, Nigeria, or Mozambique while the mean of the second tercile is 39, which is the level of state capability of Mexico, Sri Lanka or the Philippines. In GDPPC the mean of the first tercile is P$3340, around Zambia whereas mean of the second tercile is P$13,476, around Colombia's level.

Moving from the mean of the first tercile of countries to the mean of the second tercile in GDPPC produces more gains than the similar move of state capability for Basic Needs and Foundations of Wellbeing (and SPI) but is less important for Opportunity. Fig. 7 also shows, similarly for Figs. 5 and 6) that for SPI and its three components the importance of GDPPC rises and then falls. The coefficient on WGI SC is linear but the importance of moving across the terciles rises for each indicator because the gaps are larger as the mean of the second tercile is 39 (20 points higher than the mean of the first tercile) whereas the mean of the highest tercile is 71, 32 points higher.

Fig. 8 extends Fig. 4 (showing the total gain from gains in NDI) and Fig. 7 (showing the decomposition across the three components of NDI) to show the relative contributions of growth (GDPPC) and governance. In Fig. 8 we add the impact of WGI SC and POLITY(K) to get a total gain from governance. Fig. 8 shows these gains for SPI, its three components and its 12 subcomponents. This decomposes the total gain.

First, Fig. 8 shows the large gains from increased NDI shown in Fig. 4 for the three elements of Basic Human Needs–Shelter (B:HS), Water and Sanitation (B:WS), and Nutrition and Basic Health (B:NB)—come almost entirely from gains in GDPPC. For instance, the gains to Nutrition and Basic Health (B:NB) are 31.6 from GDPPC and 5.7 from Governance so the ration of growth to governance gains is 5.6.

The second finding is gains to Access to Knowledge (F:BK) and Access to Advanced Education (O:AE) over this range of development are also primarily driven by GDPPC, but with an important contribution of governance (the ratios of growth to governance gains are (3.8 and 2.4, respectively).

Third, there are three indicators where growth and governance both play important and roughly equal roles: Access to Information (F:IC), Health and Wellness (F:HW), and Personal Freedom (O:FP).

Finally, we see that the reason for the relatively small contribution of national development to progress on the four indicators in Fig. 4 above is that for these four the contribution of growth is either zero Personal Safety (B:SF)) or negative–Political Rights (O:RP), Inclusiveness (O:IV), and Environmental quality (F:EQ) and this loss from growth is offset by the positive gain from governance.

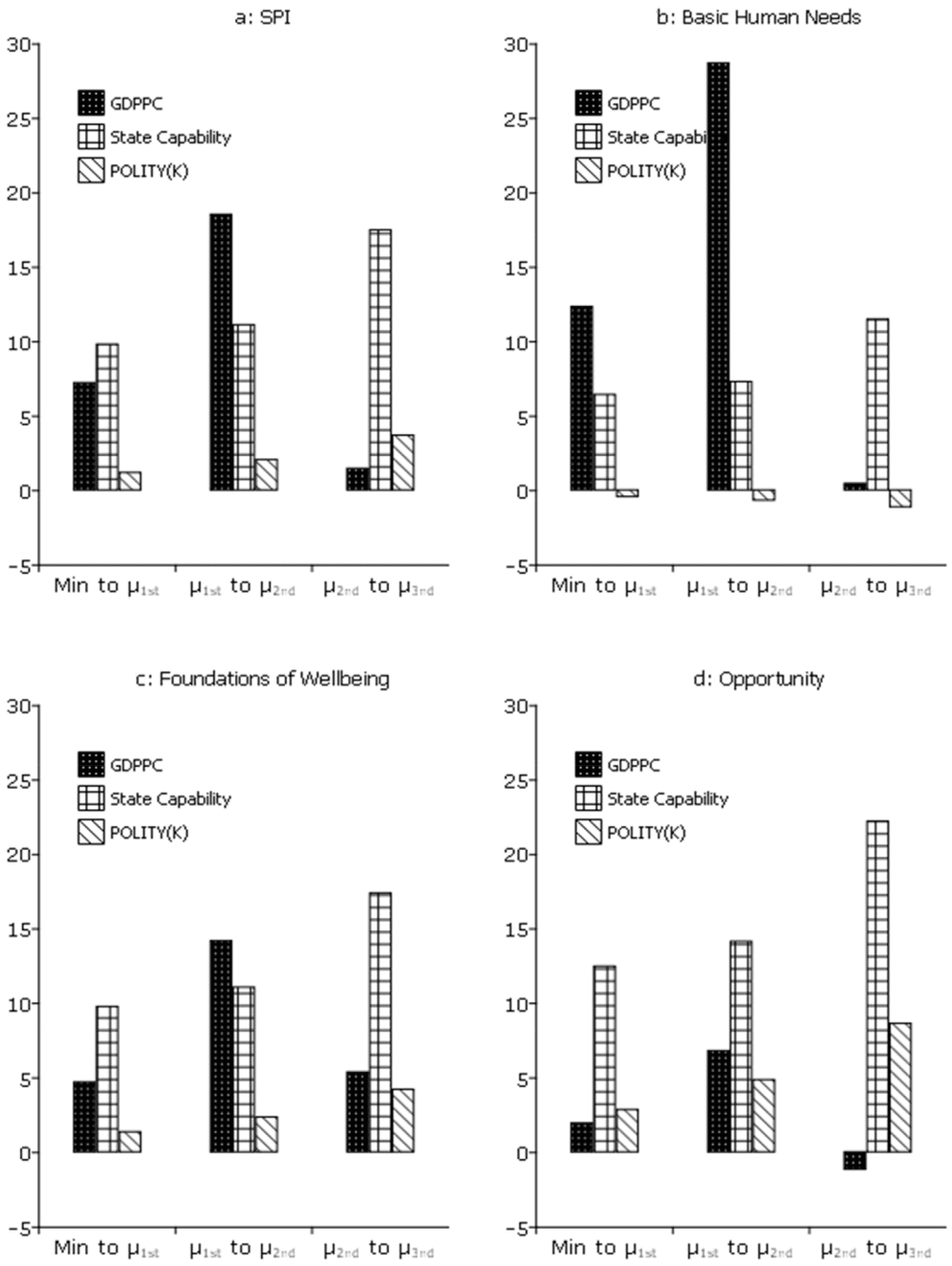

Fig. 7. a: SPI Fig. 7b: Basic Human Needs c: Foundations of Wellbeing Fig. 7d: Opportunity. Source: Author's calculations.

5. Why the strong connection between human wellbeing and national development (including economic growth) makes sense—and nothing else really does

Suppose you accept the argument that there are physical outcomes that nearly everyone wants–adequate nourishment, access to safe water, their children to survive, a quality education for their child, access to electricity–and hence these are useful measures of the wellbeing across countries. I argue that it is nearly impossible to also believe that these indicators are not very tightly linked with national development and, within that, the level of GDPPC, for three reasons.

First, the variation across countries in GDPPC is massive so that any connection between these physical indicators of wellbeing and GDPPC has the potential to explain large variations in those indicators.

Second, while GDPPC is not an indicator of wellbeing, higher levels of GDPPC nearly always translate into *both* higher levels of consumption expenditures of the typical household *and* higher levels of government spending.

Third, the very belief that an indicator is an important element of wellbeing implies that is should be responsive to consumption and government expenditure—and to a capable and responsive state–and moreover, that it should be *inelastic* or *unresponsive* to other factors, like differences in relative prices.

5.1. *Current differences across countries in GDPPC and state capability are massive*

It is simple empirics that the larger the variation in an "explanatory" variable X the more precision in estimation and the larger the potential the variable has for "explaining" (in a proximate sense) variation in

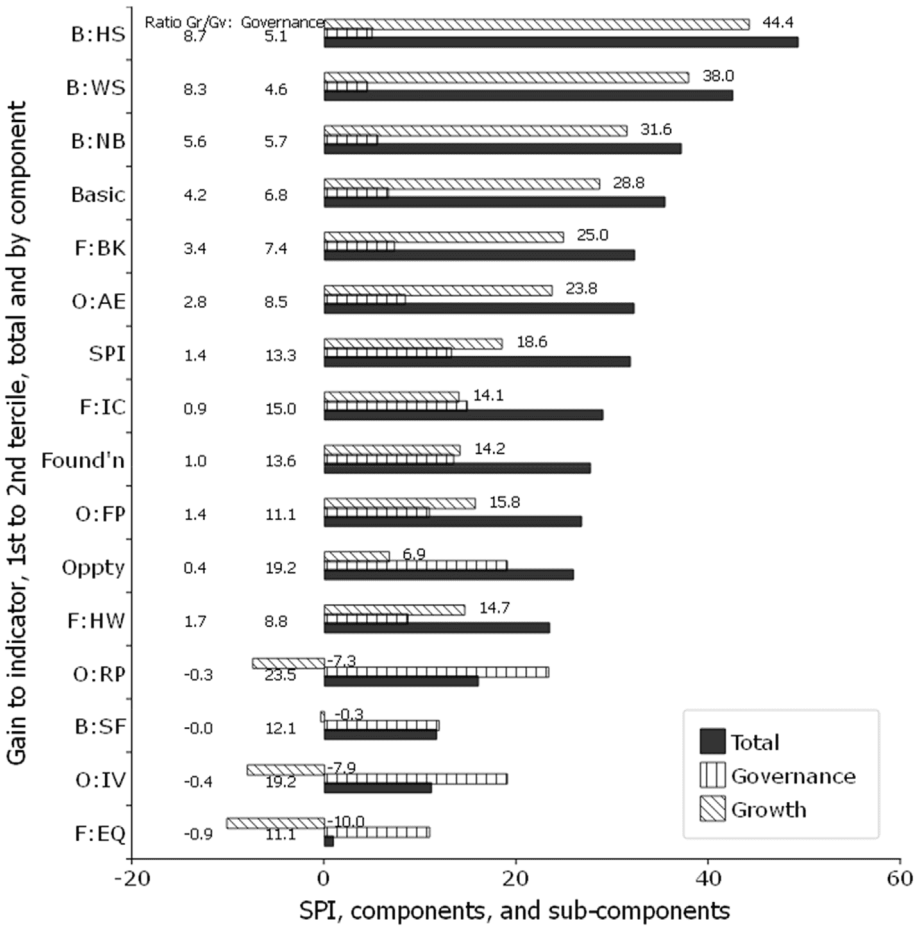

Fig. 8. Most of the improvement in wellbeing indicators measuring basics comes from growth, most of the gains in opportunity come from state capability. Source: Author's calculations.

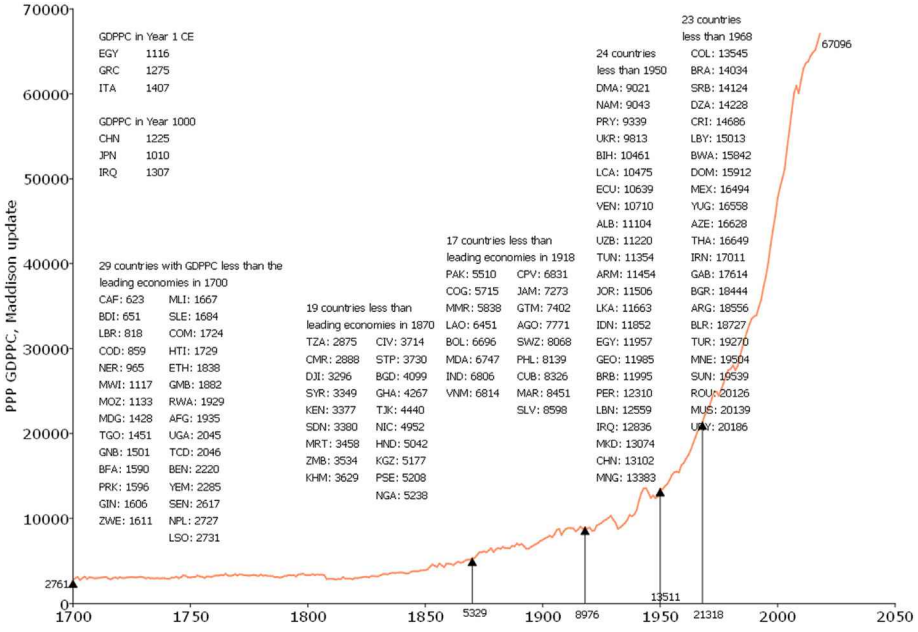

Fig. 9. The differences in GDPPC across countries in the world today (2018) is absolutely larger than it has even been—and spans the range of human history. Source: Author's calculations with updated Maddison data Maddison Project Database, version 2020. Bolt and van Zanden (2020).

outcomes. In absolute terms the cross-national variation in GDPPC is as large as it has even been in human history. The range of GDPPC across (nearly all) recorded human history is represented across countries today. Fig. 9 uses the updated and rescaled Maddison data set (Bolt and van Zanden, 2020) to show the historical evolution since 1700 of the mean GDPPC of the three leading economies in the world at the time, arrayed against the 2018 GDPPC of countries today.

There are 29 countries with 2018 GDPPC less than the leading economies had in 1700 (M$2761).[14] In fact, many of those countries have GDPPC only modestly higher (less than twice) the GDPPC of leading economies in year 1 CE. In addition, there are another 19 countries with GDPPC lower than the leading economies in 1870 and another 17 with GDPPC less than the leading economies in 1918 (100 years before the latest data in 2018). Of the 112 "developing" economies the median country (which is Vietnam) has GDPPC of M$6814 which is a factor of 10 lower than that of the leading economies of M$67,086 and the 25th percentile developing country (Nepal) is at M$2727 a factor of 25 behind the leading countries.

If there is an empirical connection between a wellbeing outcome and GDPPC the massive variation in GDPPC makes it possible to explain very large parts of the variation in the outcome.

The same point true of state capability as the difference in state capability in 2018 between the top 5 (Norway, Netherlands, Finland, Switzerland, and Singapore) and the bottom five countries (Central African Republic, Democratic Republic of Congo, Haiti, Sudan, Iraq) is massive.

5.2. *Differences in GDPPC produce expanded consumption and government spending*

National accounting can be based on "sources" or "uses" and the identity for uses is that national product is used in consumption, government spending, investment (both private and public) and net exports (exports less imports).

$$GDP \equiv C + I + G + (X - M)$$

Given the differences in levels of GDP per capita across countries are so massive, nearly all the variation across countries in the per capita consumption and government spending per capita are driven by differences in GDPPC as, while countries may differ in the proportion that is C or G but these differences are bounded (and small) compared to variation in GDPC.

Table 3 shows the level of national accounts consumption and of government spending per person in PPP from the PWT 10.0 data for quartiles of country GDPPC and for selected countries, contrasted with estimates of US spending on specific categories.

Suppose outcomes wellbeing are driven by some combination of private good and public goods. For example, perhaps nutrition outcomes are a function both of food intake and local prevalence of diseases that inhibit food intake leading to nourishment (e.g. diarrheal diseases, helminths) therefore both private and public spending matters. The availability of safe water can be created through a combination of private expenditures (e.g. chlorine treatments) or public expenditure (e.g. safe water from municipal infrastructure). Then, to the extent that higher economic productivity leads to greater availability of consumption in the hands of households who allocate it to their priority uses and greater government spending (whose efficacy at raising priority outcomes is mediated by how capable and responsive the state is) then this is likely to be associated with better outcomes.

The stark fact of Table 3 is that countries with low levels of GDPPC just have very low levels of consumption per capita and government spending per capita. This is adjusted for purchasing power and hence takes into account that government services are much cheaper in poor countries (e.g. the price of G in Ethiopia is four times lower than in the USA so it takes four times as much spending at official exchanges rates in the USA to produce the same "quantity" of G as in Ethiopia).

Table 3
Estimates of national accounts consumption expenditures and government, 2017

country/aggregate	Consumption expenditures, 2017 (in PPP)	Government, 2017 (in PPP)
DRC	$844	$64
Ethiopia	$1303	$255
Average, quartile 1 GDPPC	**$1746**	**$322**
Pakistan	$3779	$430
India	$3851	$432
Nigeria	$4283	$312
Average, quartile 2 GDPPC	**$5301**	**$1675**
China	$5533	$1941
Indonesia	$5889	$1147
Egypt	$8154	$1194
Brazil	$9082	$2691
Average, quartile 3 GDPPC	**$11,274**	**$4348**
Mexico	$12,446	$3239
Malaysia	$15,885	$3358
Average, quartile 4 GDPPC	**$23,584**	**$9032**
Germany	$26,949	$9017
USA	$42,786	$7026
Per person spending in the USA, specific categories		
Food	$5501	
Entertainment	$2281	
Cellular phone service	$797	
Pets	$505	

Sources. Penn World Tables 10.0 (Feenstra et al., 2015).
US consumption expenditures by category are based on multiplying the PWT10.0 estimate of consumption expenditures by the estimated consumption share from US Consumer Expenditure Survey 2017, Table 1203.

The typical country in the bottom quintile has just P$322 per person in G to devote to all uses: education, health, infrastructure (roads, power, water, sanitation), law and order, justice, regulation (safety, environment, economic). A stark comparison the typical person in the USA spends *twice as much* on their cell phone services as these governments can devote to *all* uses.

One could imagine raising wellbeing by allocating from private consumption expenditures to more government to create more and more effective public spending. But private spending per person across all uses is only P$1746. As we see below, most of that private spending by the typical (median) household in a poor country is spent on necessities—like food. Moreover, even if 100 percent of the consumption spending of households in the poorest quartile went to food they would still spend *less than a third* of what the average US person spends on food. So, unlike in a rich country, reallocating from C to G in a poor country is exchanging necessary private spending for necessary government spending.

The average government spending in the second quartile of counties by GDPPC is much higher, but still only P$1675 per person. Even if that amount is spent according to the all and only highest priorities for raising wellbeing and even if that amount is spent with very high efficacy, it still is a limited amount to address all of the pressing needs a government would like to be able to meet. Even these governments spend less on all uses than the typical US person spends on entertainment alone.

As with the relationship between wellbeing and national development, there is also a very tight relationship between G and GDPPC. One, the elasticity of G wrt GDPPC is greater than 1 (our estimate is 1.12) which implies that G rises more than proportionately with GDPPC so more growth tends to lead to not just more G but more G as a share of GDP. Second, the relationship is very tight (R2 of 0.913) so there is a "empirically necessary and sufficient" relationship between government spending and GDPPC in that "no government has G per capita higher than

[14] "M$" is Maddison dollars, which are PPP adjusted but not directly comparable to the PWT PPP estimates.

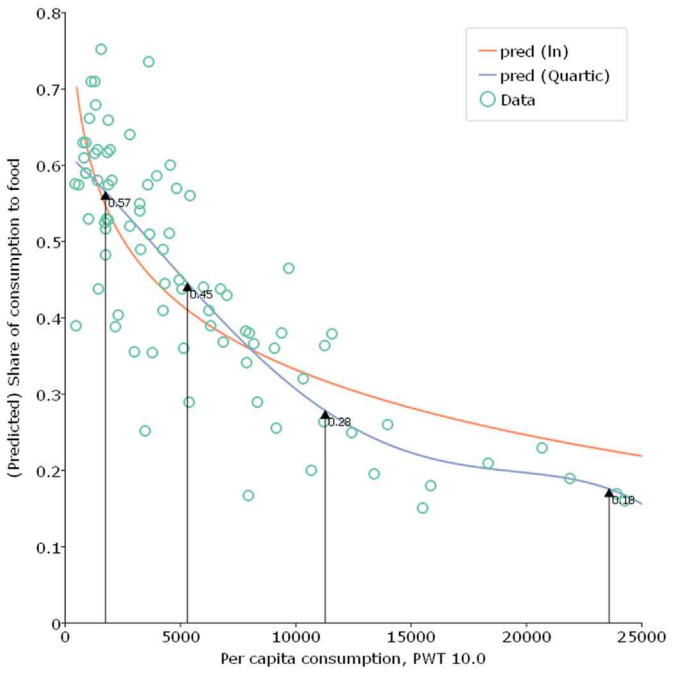

Fig. 10. The share of household expenditures going to necessities (here, food) falls as overall consumption expands. Sources: PWT 10.0 (Feenstra et al., 2015) and Pritchett and Spivack (2013) for data on food shares.

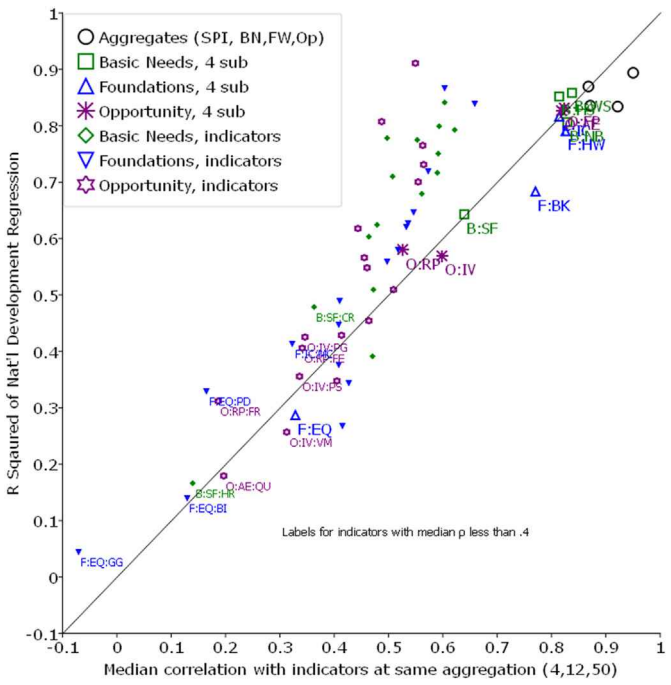

Fig. 11. The strong connection between wellbeing measures that are strongly correlated with other and also correlated strongly correlated with national development. Source: Author's calculations.

X without GDPPC higher than Y."

This point is important as sometimes advocacy for a greater emphasis on economic growth is mistakenly taken as a prioritization of "private" over "public" goods whereas the only way governments can adequately fund public goods is to be able to tax a productive economy and total tax revenues tend to grow *more than proportionally* with economic growth.

5.3. At low levels of income necessities have a high budget share and low price elasticity (and hence the relationship with consumption is tight)—and how exceptions illustrate the rule

A working definition of a "necessity" is something for which the marginal utility (incremental benefit per unit of consumption) goes to infinity as consumption nears zero but then declines (perhaps rapidly) as consumption increases. If I am suffocating from a lack of oxygen more oxygen is worth life itself but adding oxygen to normal room air produces little or no gain. This simple mechanism produces the "diamond-water paradox" and the Engel curve. This implies that at very low budgets nearly all of the budget will be devoted to necessities and the marginal propensity to consume necessities will be high (but falling).

Fig. 10 shows estimated Engel curves across countries (one standard, food share on natural log and one food share with quartic terms) relating the budget share of the median household to the per capita national accounts consumption.[15] At the mean of the bottom quartile of countries the predicted food share is 57 percent, falling to 45 percent at the average consumption per capita of the second quartile of countries by GDPPC and then falls to 18 percent for the richest quartile of countries. Of course, not all food expenditure is "necessary" and not all necessities are food, but the basic point is validated by any empirical approach to examining the connection between the average budget shares and marginal propensity to spend on necessities: it will be a very large share at low incomes and decline to a very small share at high incomes.

The second point, about price elasticities, is less appreciated. We should expect *outcome* measures on important elements of well-being—necessities–to be very inelastic with respect to anything but income, and in particular are likely to be price inelastic.

This implies the basic logic of the argument that growth will not reliably produce gains in important and universally shared direct physical indicators of wellbeing—like nutrition, basic education, health, water and sanitation—has it *exactly backwards.* These elements of well-being that are "basic human needs" should have the *strongest* relationship to gains in consumption from low levels precisely because they are so important that (i) incremental budget shares to necessities will be high and (ii) low "price" elasticities will translate into large variations in the sacrifice made to achieve gains, not as much in the actually indicators themselves.

One comparison that illustrates this is to compare for each component and indicator the R2 explained by national development and the correlation of that component or indicator with all other components or indicators. We should expect "necessities" to have very high correlation *both* with national development *and* with each other as the analog of the "budget expansion path"–the "national development expansion path"—should look similar for all necessities. That is, on the simple conjecture that budget shares are high and price elasticities low for necessities they should be highly correlated with each other and highly correlated with national development.

Fig. 11 shows the scatter plot of the R2 of regressing each component or indicator on the three national development variables (on the y axis) and the median correlation of each component or indicator with all others at its same level of aggregation (e.g. the four (SPI and its three components), the 12 sub-components, and the 50 individual indicators. Table 4 shows the R2 and median correlation of the 50 individual indicators for the 10 highest correlation and the 10 lowest correlation indicators.

There are three facts illustrated in Fig. 11 and Table 4.

First, as we have seen above, the four main aggregates, by their nature

[15] The data on food share from household consumption surveys is from various sources for various years (see Spivack and Pritchett, 2013 for details). I am not describing this particular regression in much detail as (i) the point is mainly illustrative of a well-known fact and (ii) the actual quantitative parameters of the Engel curve relationship are very robust across space and time and this estimated curve is pretty much exactly like all the others (Spivack and Pritchett, 2013).

Table 4
The indicators of wellbeing with the highest and lowest correlations with other indicators and R2 with national development.

Code	Name/description	R2 of national development (SC, PK, quartic GDPPC)	Median correlation with all other indicators	Average
Ten (of 50) indicators with the highest median correlation with other indicators				
F:HW:AE	Access to essential health services	0.839	0.659	0.749
F:IC:IP	Internet penetration	0.866	0.603	0.735
O:FP:CP	Corruption	0.912	0.549	0.730
B:WS:US	Population using unsafe or unimproved sanitation	0.841	0.604	0.722
B:HS:IA	Household deaths attributable to indoor air pollution	0.792	0.622	0.707
B:WS:UW	Population using unsafe or unimproved water sources	0.799	0.593	0.696
B:NB:CS	Child stunting	0.751	0.591	0.671
B:HS:CF	Usage of clean fuels and technology for cooking	0.776	0.553	0.664
O:FP:VE	Vulnerable employment (contributing family workers and own-account workers as % of total employment)	0.765	0.563	0.664
B:NB:CM	Child mortality	0.716	0.591	0.654
Ten (of 50) indicators with lowest median correlation with other indicators				
F:IC:MC	Media censorship	0.414	0.323	0.369
O:IV:PS	Equality of political power by social groups	0.356	0.337	0.346
F:BK:PE	Primary enrollment	0.268	0.415	0.342
O:IV:VM	Discrimination and violence against minorities	0.258	0.313	0.286
O:RP:FR	Freedom of religion	0.312	0.187	0.249
F:EQ:PD	Outdoor air pollution attributable deaths	0.329	0.165	0.247
O:AE:QU	Quality weighted universities	0.179	0.196	0.187
B:SF:HR	Homicide rate	0.167	0.140	0.153
F:EQ:BI	Biome protection	0.140	0.129	0.135
F:EQ:GG	Greenhouse gas emissions	0.044	−0.071	−0.013

Source: Author's calculations with Social Progress Index data.

of being the sum of lots of correlated variables, are very highly correlated with each other and with national development. Second, most of the 12 sub-components are also very highly correlated with each other and with national development. Seven of the 12 indicators have both median correlations with each other and R2 above (about) 0.8: three of the four components of Basic Needs (Nutrition and Basic Health, Water and Sanitation, and Shelter), two components of Foundations of Wellbeing (Health and Wellness, Access to Information and Communication) and two components of Opportunity (Personal Freedom and Choice and Access to Advanced Education). Access to Basic Knowledge does not quite meet the 0.8 threshold but also has a high median correlation (0.771) and relatively high R2 (0.683). Pritchett and Lewis (2021) use these correlations amongst indicators to define a measure of basics.

Second, Fig. 11 and Table 4 shows that there are a set of individual indicators that are strongly correlated both with other indicators and strongly associated with national development. A high median correlation suggests that a country is unlikely to have high levels of wellbeing on many other indicators and not make significant progress on these indicators. The top 10 highest average of median correlation and R2 are (nearly all) indicators one could call "necessities": child mortality, child stunting, safe/improved water, safe/improved sanitation, using improved cooking fuels, fewer deaths from indoor air pollution, access to essential health services. The two slightly puzzling entries are internet penetration and corruption (the latter puzzling because it isn't a direct household consumption item).

Third, there are also clearly a set of specific indicators (not aggregates) that have both low correlation with other indicators and are not highly correlated with national development. These are the exceptions to a general claim that "national development solves all ills." These exceptions are themselves illustrative and fall into three types.

The first type are indicators that may not command powerful action even in a rich country with a capable and responsive state because their benefits fall on a minority—and perhaps on minorities that are unpopular or marginalized: equality of political power across social groups (O:IV:PS), freedom of religion (O:RP:FR), and discrimination and violence against minorities (O:IV:VM). This doesn't make these any less crucial as features of a just and fair society but one can understand that a country could have high levels of material wellbeing (and have met necessities for nearly all the population) and have a capable and democratic state and still face serious issues with discrimination against minorities (my native country, the USA, being an example).

The second type is environmental indicators, which themselves fall into two categories.

One type of environmental indicators is those for which the harms (scope of the negative externality) are (almost all) within the country. Outdoor air pollution attributable deaths (F:EQ:PD) is the result of negative environmental externalities that require high levels of state capability to regulate but which also, if unregulated, tend to grow with economic growth so that there is a non-linear relationship with growth (often called an "environmental Kuznets curve" (Grossman and Krueger, 1995; Dasgupta et al., 2002)) and hence the correlation with other indicators and national development is not simple. This is even though at very high levels of development a political "internalization of the externality" will almost certainly bring this type of environmental bad under control.

The other type of environmental indicators is those whose impact is either completely or totally global. The obvious example is greenhouse gas emissions (F:EQ:GG) and biome protection (F:EQ:BI) where a major challenge is precisely that the geographic scope of the externality is global so countries do not bear the full costs of their emissions or of the loss in biodiversity they might cause.

The third type of indicators are those that are apparently subject to very specific determinants that go beyond national development or that for which the particular empirical strategy fails. For instance, many of these variables in their raw form are very highly skewed. The homicide rate (B:SF:HR) and quality weighted universities (O:AE:QU) and primary education (F:BK:PE) are all examples where most countries have very nearly the same value and only a few countries are very different and hence the absolute differences among countries with nearly the same ranking will be very small and this skewed distribution with a large cluster and a few outliers makes it hard for regressions to work well.

In summary, for "necessities" that matter broadly for human wellbeing we should expect not only that national development matters but we should not be surprised that national development is essentially *all that matters.* This is because national development is a means of nominating and solving priority problems and hence if one has greater income (that can be used both privately and expands resources in the public sector), a capable state, and a responsive government it would be surprising, indeed astounding, if common, shared, priority problems did not improve.

5.4. *Why I use levels*

Everything I have said so far about "growth" is based on empirical

associations between *levels* of income and *levels* of wellbeing indicators. People often try and refute the strong and tight connection between growth and wellbeing by using data of *changes on changes* rather than data on levels. This leads to three common econometric mistakes.

First, this process of using "changes on changes" can transform a strong result into a low powered failure to reject simply by throwing away most of the available variation in the independent variable. The entirely predictable consequence of using "changes on changes" both increases the attenuation bias from measurement error, which depends on the ratio of noise to signal and hence is worsened with reduced signal and raises coefficient standard errors.[16]

Second, "changes on changes" regressions nearly always mis-specify the dynamics. Suppose there is a stable long-run relationship between income and health and an econometric study uses only changes in income and changes in population health over five-year periods. Those changes would have two components: (i) the move along the long-run stable relationship and (ii) the adjustment dynamics towards the long-run relationship from any point off the equilibrium relationship. Often when studies find that there is no relationship in the changes-on-changes data it is because they have not included any adjustment dynamics and the inclusion of those dynamics actually produces from short run data estimates of the long-run impact that are exactly the same as the levels on levels estimates (see Pritchett and Viarengo, 2010 documenting this methodological point for cross-national life expectancy regressions).

Third, if the long-run relationship is non-linear then using changes on changes often loses the ability to estimate that relationship. Suppose the long-run relationship between Y and X was an "S" curve in levels. We can identify that in levels data if we have countries at all parts of the relationship. However, in "changes on changes" data each country is only moving over a quite small part of the S curve and without an complex interactive specification in levels and changes the econometrics with only "changes on changes" could never estimate the true S shape.

6. Conclusions and implications

There is a joke among econometricians that if one tortures the data long enough it will confess. I disapprove of torture, even of data. I find that if you ask the data politely, are patient, and provide flexibility in what the data can say, it will talk. The data on national development with data on human wellbeing say three important things.

First, high levels of national development are strongly empirically necessary and sufficient for high levels of overall human wellbeing. There are no countries with unproductive economies (low GDPPC), weak administration (low WGI SC) and unresponsive governments (low POLITY(K) that achieve even above average human wellbeing in the Social Progress Index or any of its three components (Basic Human Needs, Foundations of Wellbeing or Opportunity). While there is some scope for most countries to achieve higher levels of wellbeing at any given level of national development, these gains are quite limited relative to the upside potential of national development.

Second, the absolute and relative importance of economic growth for improving human wellbeing depends on a country's current level of GDPPC. Growth is much more important for raising aggregate measures of well being in poorer than in richer countries.

Third, whether "growth" or "governance" matters most depends on the economic characteristics of that indicator. For some indicators, for which the actions households can take with their income are essential to improvements, like nutrition and basic health outcomes, water and sanitation, shelter, economic growth is far and away the most important element of national development. On the other hand, for indicators that require effective public action—like personal safety and political equality and environmental quality—governance matter most for improvements.

I realize that a paper using cross-national data and simple OLS seems old-fashioned and that the findings may seem too commonplace to justify an article, but I think there are three important debates that need to be grounded in these facts.

First, many people and politicians from OECD countries are pushing back against the concept of economic growth generally and against the measure of GDP as a useful and reliable measure of economic productivity. The sharp non-linearity of the impact of growth and the importance of growth for accomplishment of basics is a useful reminder there can be no "global" conversation about the importance of growth for wellbeing. Most of current developing world is in the range of GDPPC in which the data say economic growth is at its *most* important for raising human wellbeing (especially on basic needs). Given the massively different levels of GDPPC even if everyone in two different countries had exactly the same preferences and even if their governments perfectly reflected those preferences, governments could have very different *priorities* if they are staring from different levels of GDPPC. There is no contradiction, or even surprise, if New Zealand wants to downplay economic growth and yet countries and governments in India, China, Pakistan, Nigeria, Ethiopia, Egypt governments, want to remain focused on growth. This makes it particularly important that development actors and agencies based in and/or influenced by rich countries recognize that growth is a development priority even it isn't a priority for their country.

Second, there has been a major movement within development economics to try and move research and policy focus away from national development towards "small" issues of program and project design, with the notion being that "rigorous" evidence about "what works" for projects and programs can lead to major improvements in outcomes (Banerjee and Duflo, 2011). There is no question that more effectively designed and implemented programs would lead to better outcomes in a variety of domains (from poverty programs to education to health to microfinance, etc.) However, this tactic (and particularly the importance of "rigorous" evidence) has been vastly oversold. For instance, a recent op-ed from JPAL titled "growth is not enough" asserted that:

But for millions of people living in poverty, growth is not enough. Specific, targeted social programs based on rigorous empirical evidence are equally important to prevent people from being left behind.

That social programs are "equally important" for explaining the level or decline in poverty is not even within an order of magnitude of correct (Pritchett, 2020). This paper shows that the same is roughly true for many other indicators of wellbeing. Once one has conditioned on just three components of national development there is only 10 to 15 percent of the total variation in wellbeing outcomes to be explained and "specific, targeted, social" programs as a causal factor in dependent of national development are just one possible element that could explain the remaining variation—but there is no evidence that these types of interventions have been even 1 percent of the explanation of cross-national differences in wellbeing outcomes. The limited upside scope for those types of improvements at a given level of national development needs to be factored into the formation of a balanced strategy for research and development that, at the very least, has a research and practical strategy for broad-based economic growth and for generally improved state capability over and above the attention to programmatic design issues.

Third, there are powerful forces pushing for "growth denialism"—the notion that high levels of human wellbeing can be achieved in the absence of economic growth. One path to "sustainability" would be to curtail economic growth that uses natural resources, creates pollution, or threatens climate stability. Perhaps this is the most sensible policy for the now rich countries. However, there is no evidence that it is possible to reach high levels of human wellbeing, or even, for that matter, universal access to basics like adequate nutrition, access to water and sanitation, electricity, basic health, for many countries without national development, including sustained economic growth. A position in favor of

[16] This is also of course a problem with techniques like instrumental variables that trade off consistency for efficiency as one can move from a "rejection" with OLS estimates to a *larger magnitude* coefficient estimated with IV but the larger standard errors from weak instruments also increase the standard errors so much so that one "fails to reject" a null of a zero coefficient.

sustained and sustainable growth is not a denial of the enormous environmental stresses and potential environmental "bads" through wastes, pollution of air, water, and soil, and risk to climate sustainable that growth might create but simply an acknowledgement of the importance of improving human wellbeing both now and in the future, especially for those who currently lack even the basics.

Declaration of competing interest

I, the author of this manuscript submitted, have no conflicts of interest of any type.

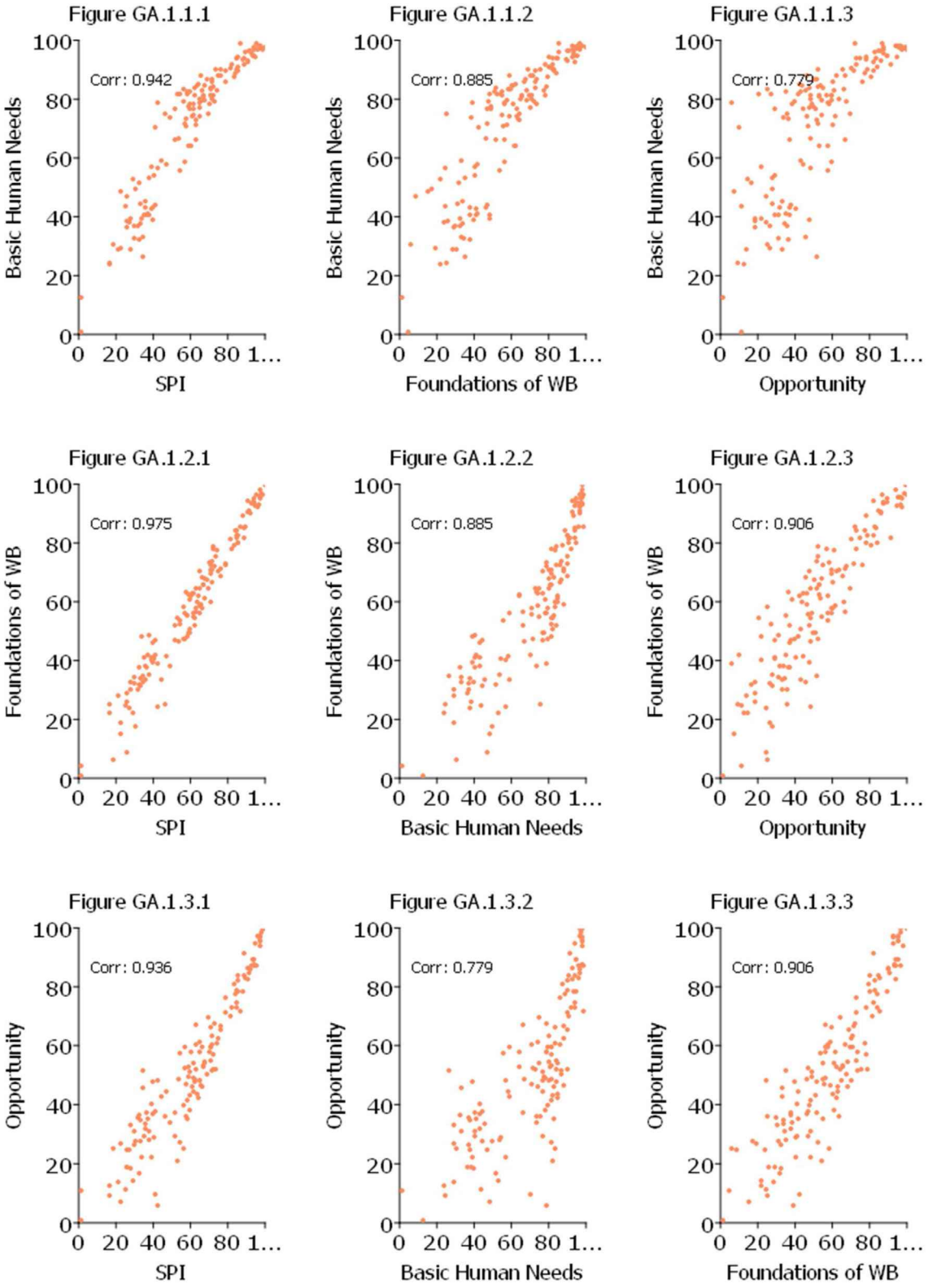

Fig. GA.1.

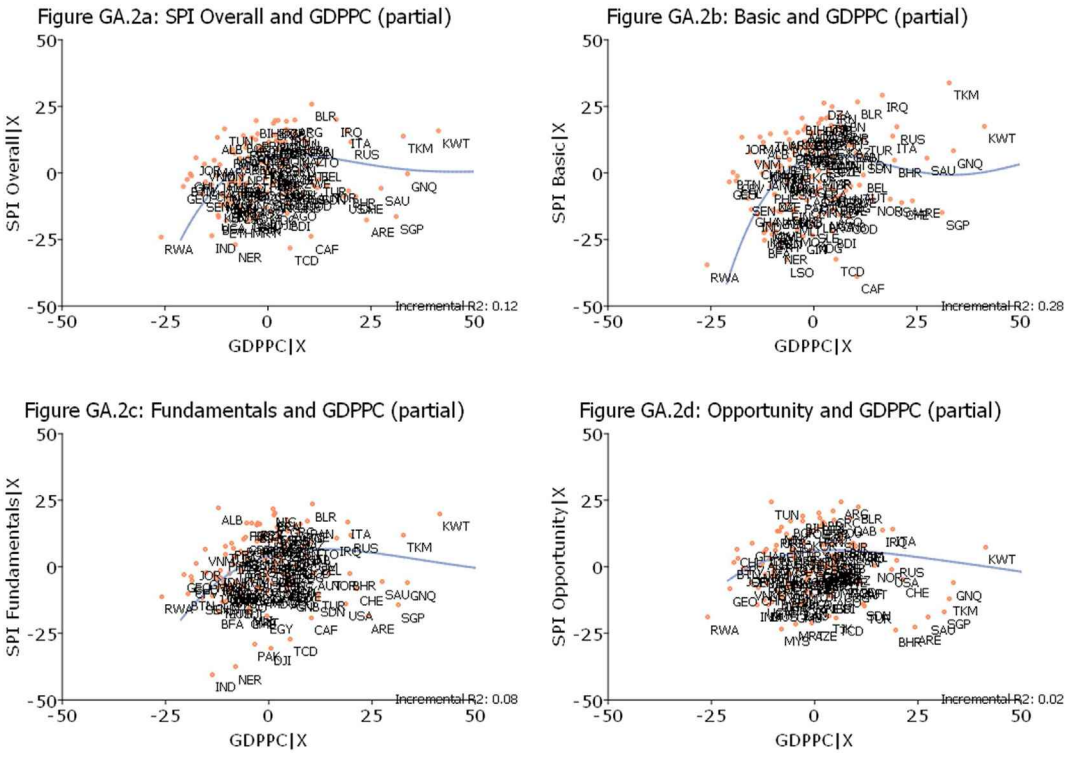

Fig. GA.2. a: SPI Overall and GDPPC (partial) b Basic and GDPPC (partial) c Fundamentals and GDPPC (partial) d Opportunity and GDPPC (partial).

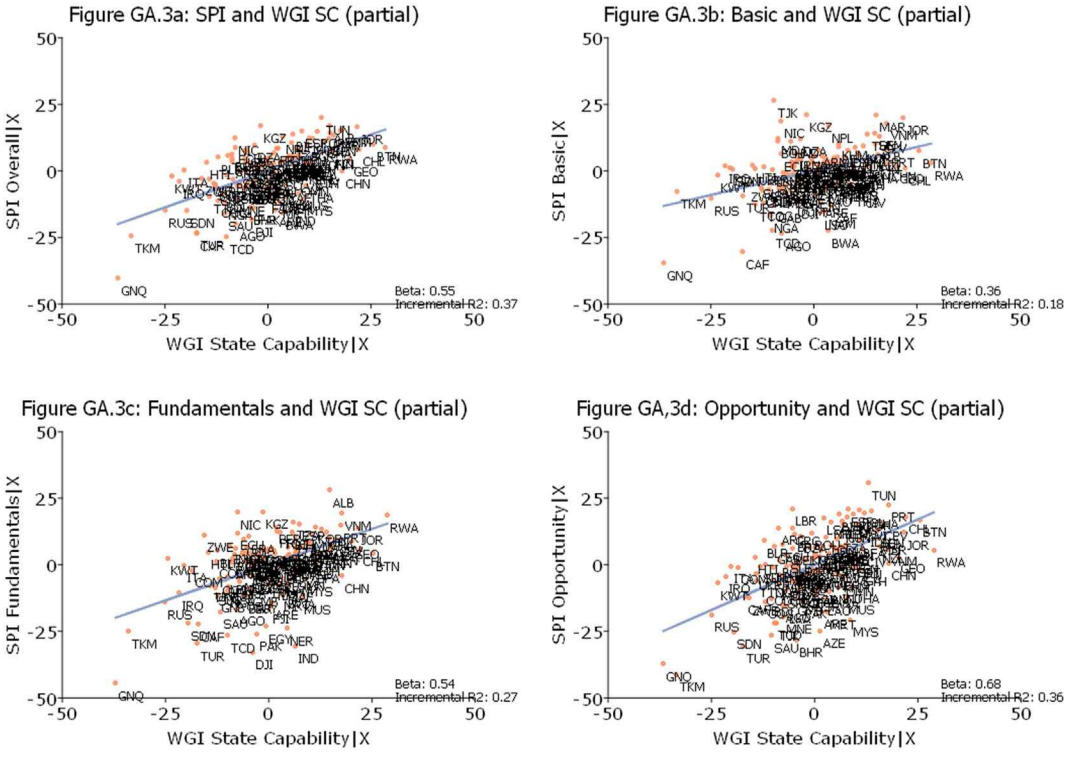

Fig. GA.3. a: SPI and WGI SC (partial) b Basic and WGI SC (partial) c Fundamentals and WGI SC (partial) d OPPortunity and WGI SC (partial).

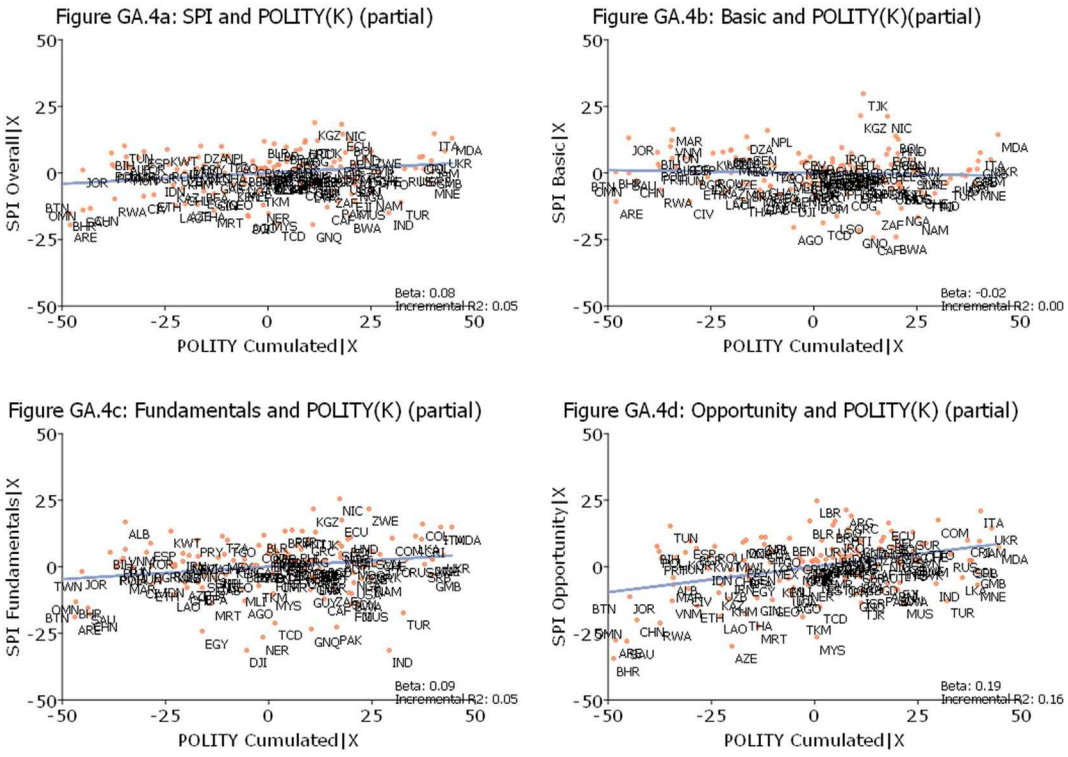

Fig. GA.4. a: SPI and POLITY (K) (partial) b Basic and POLITY (K) (partial) c Fundamentals and POLITY (K) (partial) d OPPortunity and POLITY (K) (partial).

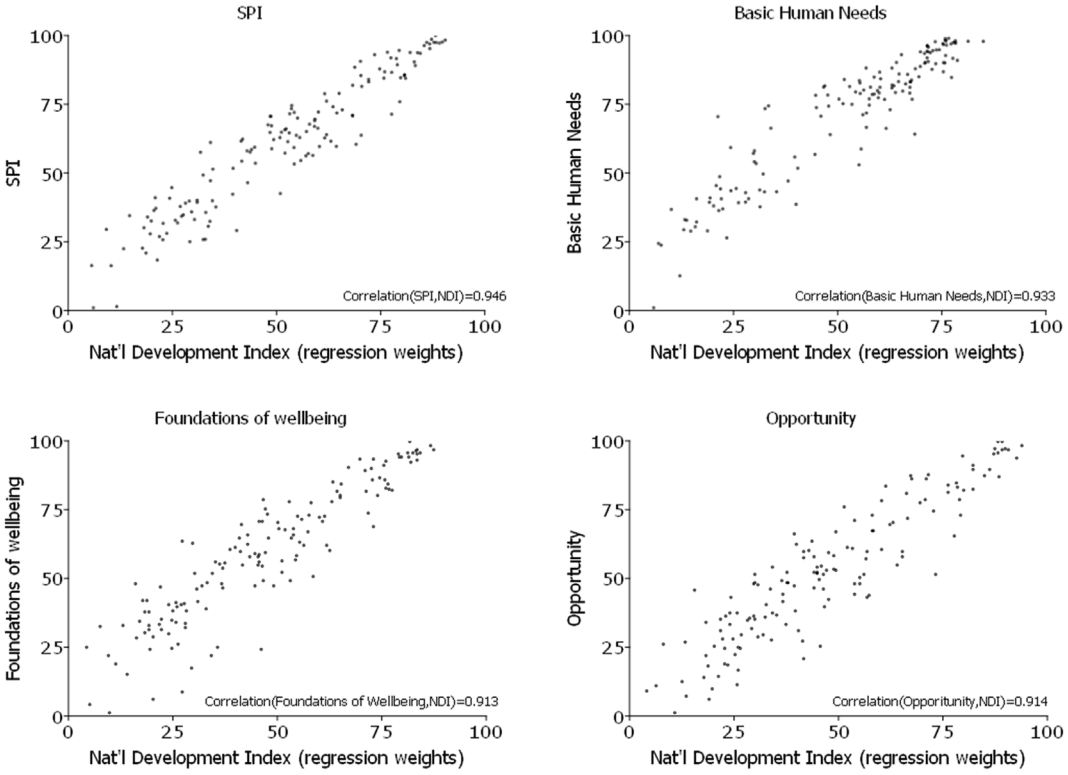

Fig. GA.5. Scatter plots of SPI, Basic Needs, Foundations of Wellbeing, and Opportunity with NDI.

References

Andrews, Matthew, Pritchett, Lant, Woolcock, Michael, 2016. Building State Capability: Evidence, Analysis, Action. Oxford Univerity Press, Oxford, UK.

Banerjee, Abhijit, Duflo, Esther, 2011. Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. Public Affairs, New York.

Bolt, Jutta, van Zanden, Jan Lutten, 2020. Maddison Style Estimates of the Evolution of the World Economy. A New 2020 Update, vol. 15. Maddison-Project Working Paper, WP.

Camerer, Colin F., Anna, Dreber, Holzmeister, Felix, Ho, Teck-Hua, Huber, Jürgen, Johannesson, Magnus, Kirchler, Michael, Nave, Gideon, Nosek, Brian A., Pfeiffer, Thomas, et al., 2018. Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. Nat. Hum. Behav. 2 (9), 637–644.

Clemens, Michael, Hamilton, Kirk, 1999. Genuine savings rates in developing countries. World Bank Econ. Rev. 13, 333–356.

Dasgupta, S., Laplante, B., Wang, H., Wheeler, D., 2002. Confronting the environmental Kuznets curve. J. Econ. Perspect. 16 (1), 147–168.

Dollar, David, Kleinenberg, Tatjana, Kraay, Aart, 2016. Growth is still good for the poor. Eur. Econ. Rev. 81, 68–85.

Drumm, Brian Richardson, 2015. Distinguishing earth, water, fire, and air: factor Analysis to determine the four fundamental elements of state capability. Bachelor's Thesis. Appl. Math. Harv. Coll.

Easterly, William, 1999. Life during growth. J. Econ. Growth 4 (3), 239–276.

Feenstra, Robert C., Robert Inklaar, Timmer, Marcel P., 2015. The next generation of the Penn world table. Am. Econ. Rev. 105 (10), 3150–3182.

Filmer, Deon, Pritchett, Lant, 1999. The impact of public spending on health: does money matter? Soc. Sci. Med. 49 (10), 1309–1323.

Goldstone, Jack, Ulfelder, Jay, 2004. How to construct stable democracies. Wash. Q. 28, 7–20.

Grossman, Gene M., Krueger, Alan B., 1995. Economic growth and the environment. Q. J. Econ. 110 (2), 353–377.

Inglehart, Ronald F., 2008. Changing values among western publics from 1970 to 2006. W. Eur. Polit. 31 (1–2), 130–146.

Ioannidis, John P.A., 2005. Why most published research findings are false. PLoS Med. 2 (8), e124.

Kaufmann, Daniel; Kraay, Aart; Mastruzzi, M;. "Governance Matters V: Governance Indicators for 1996-2005,"

Li, Chuan Zhong, Lofgren, Karl-Gustaf, 2011. In: Green National Accounting and Sustainability. Edward Elgar Publishing, Cheltenham, UK.

McKenzie, David, 2020. If it Needs a Power Calculation, Does it Matter for Poverty Reduction?, vol. 127. World Development.

Narloch, Ulf, Kozluk, Tomasz, Lloyd, Ainsley, 2016. Measuring inclusive green growth at the country level taking stock of measurement approaches and indicators. GGKP Work. Pap. 2016 2.

Preston, Samuel, 1975. The changing relation between mortality and level of economic development. Popul. Stud. 29 (2), 231–248.

Pritchett, Lant, 2015. Can rich countries Be reliable partners for national development. Horizons 2.

Pritchett, Lant, Kenny, Charles, 2013. Promoting Millennium Development Ideals: the Risks of Defining Development Down. Center for Global Development Working Paper, p. 338.

Pritchett, Lant, Lewis, Addison, 2021. Basically, economic growth is enough, and only economic growth is enough.

Pritchett, Lant, Spivack, Marla, 2013. Estimating Income/Expenditure Differences across Populations: New Fun with Old Engel's Law. Center for Global Development Working Paper, p. 339.

Pritchett, Lant, Summers, Lawrence H., 1996. Wealthier is healthier. J. Hum. Resour. 31 (4), 841–868.

Pritchett, Lant, Viarengo, Martina, 2010. Explaining the cross-national time series variation in life expectancy: income, women's education, shifts, and what else? Hum. Dev. Res. Pap. HDRP-2010 31.

Young, Alwyn, 2019. Channeling Fisher: randomization tests and the statistical insignificance of seemingly significant experimental results. Q. J. Econ. 134 (2), 557–598.