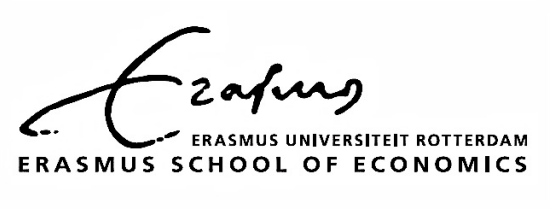

The Robustness of the Easterlin Paradox

David November
428491dn

Industrial Dynamics and Strategy Masters Thesis
Erasmus School of Economics
Supervisor: Dr. Spyridon Stavropoulos

Abstract: Easterlin's seminal paper in 1974 claimed that due to the two relationships between GDP per capita and self-reported well-being at individual and country levels, happiness was in fact stimulated by relative income. However, more recent papers have contradicted this hypothesis, claiming that his study was flawed and showing evidence of a positive correlation between happiness and absolute income. This paper aims to forward this line of thought, claiming that one of Easterlin's crucial flaws was the absence of robustness from his study, in particular, his assertion that there is a lack of correlation between happiness and GDP per capita on a country-level. We shall be reassessing this relationship, recreating his study, analysing more economic performance indicators, introducing further controls, and using a more recent dataset that includes more countries.

Contents

1. Introduction

National happiness is an increasingly prevalent topic in today's society, and in response to this, is also a highly debated topic amongst economists. Bhutan's use of a Gross National Happiness Index from 2008 is evidence that the world is moving into a more welfare-centric world with a higher emphasis on population satisfaction rather than growth and income.[1] The reason for this move can be found in a statement in the Legal Code of Bhutan that claims "if the Government cannot create happiness (dekid) for its people, there is no purpose for the Government to exist".

In the argument over the best measure of development, a further proponent of the movement away from GDP is the "Sarkozy report", a 292 page report on the measurement on social progress; a report ordered by the then French President, Nicolas Sarkozy, whose aim was to shift emphasis to the well-being of citizens and sustainability (Easterlin, 2010). In this report, the authors argue that, "emphasising well-being is important because there appears to be an increasing gap between the information contained in aggregate GDP data and what counts for common people's well-being" and as such, improvement is necessary (Stiglitz et al, 2012). There is much debate around this topic as to whether this is the correct direction in which to steer a country, as there is the potential problem that a country will miss out on potential growth and development that would benefit them more in the long run in order to pursue an inaccurate measure of their populations' happiness (Adler, 2009).

Within the GNH measure, there is still a component of income that 'contributes' to the overall happiness of people. This is a contentious issue within the field of economics; there is much debate as to whether money is what makes people happy. The debate rages on as to whether it is income (Easterlin, 1974, 1995; Veenhoven, 1988, 1991), social capital (Bjørnskov, 2003; Ram, 2010), psychological characteristics (Easterlin, 2006; Argyle, 2013), health (Palmore & Luikart, 1972), or whether another feature may be the main determinant of well-being for an individual. The importance of this debate can be seen when we consider this in line with the GNH move in Bhutan. If countries are beginning to consider the happiness of their

[1] The GNH Index is formed using 9 equally aggregated domains composed of 33 clusters, using 124 variables to yield a numerical value of satisfaction levels across the country (Ura et al, 2012).

populations as being important, then we can see these types of studies to be of major importance to policy. With little information on the topic, this is a pre-emptive solution for such a problem. Conducting a study like this can allow policy-makers to make more informed decisions, avoiding a potential misallocation of resources.

In his seminal paper, "Does economic growth improve the human lot? Some empirical evidence." Easterlin (1974) presents evidence that on an individual level, an increase in income shall have a positive impact on one's self-reported well-being, but at the same time, when looking at aggregated happiness data the difference in GDP on a country-level is limited. From these two results, he concludes that increases in absolute income have no impact on well-being, whilst it is the relative income compared to our peers that has a positive impact.

Furthermore, when he again looks at this in his 1995, in his paper, "Will raising the incomes of all increase the happiness of all?", he once again finds a similar result within countries over time. In this paper, he determines that within a country at a given time those with higher incomes are, on average, happier. However, raising the incomes of all does not increase the happiness of all. As a result, once again concludes that it is the relative level of income that contributes to happiness and not an absolute increase. Easterlin uses the phrase "well-being", where various other papers also use "happiness" or "life satisfaction", however we shall be considering these three terms to refer to the same mechanism in the literature.

In this study, we shall be questioning his results, why he sees these results, and whether his study could be improved upon. There have been multiple papers written since his original paper, questioning and refuting his claims (Stevenson & Wolfer, 2008; Veenhoven, 1991, 2013; Hagerty & Veenhoven, 2003). This has left the topic open to discussion, as there have been no concrete evidences that have solved "Easterlin's paradox". In order to do this, we shall firstly be updating his original experiment by including more countries and using a modern dataset. After this, we shall use different measures of 'development' and include further control variables in order to test the robustness of his results. The aim of this paper is not to find the focal variable that impacts individuals' happiness; the aim is instead to determine whether

economic development and income in particular have the effect on happiness as suggested by Easterlin.

The rest of the paper will continue as follows: section 2 will review the current literature on the causes of happiness. Section 3 will provide background to our data and lay out our hypotheses. Section 4 will give an analysis of the data. In section 5, there will be discussion about the evidence and the robustness of the results, and the limitations of this study. Finally, section 6 will end with my concluding remarks.

2. Literature Review

The most basic question from this topic is:

Does a higher income bring people more life satisfaction?

It seems somewhat arbitrary to say but since people spend so much time and effort in the pursuit of money that we shall assume that having higher levels of income raises overall levels of life satisfaction on average. In both of Easterlin's 1974 and 1995 papers, he indeed argues that there is a positive correlation between happiness and per capita income. However, this positive relationship is down to an increase relative to those around them rather than in an absolute sense. There is much dispute over the relativity mechanism he claims is used in this relationship, with some critical of his findings (Stevenson & Wolfer, 2008; Sacks, Stevenson & Wolfer, 2012; Ball & Chernova, 2008); and whilst others are in favour of his hypothesis (Clark, Frijters & Shields, 2008; Ferrer-i-Carbonell, 2005), they all agree on this positive correlation ceteris paribus. This also seems to be a fairly unanimous assumption within the economic community; while there is disagreement on the mechanism upon which it impacts happiness, there is broad agreement that it has at least a weak monotonic function.

The main disagreement over the mechanism is the way in which income impacts happiness. Easterlin's theory is based around James Duesenberry's (1949) 'Relative Income Hypothesis' which states that firstly an individual's utility gained from consumption and saving is a function of the consumption and savings of others around him and secondly, that this utility is also a function of past individual consumption in previous periods. From this, Abramovitz (1959) moved on to

conclude that we must be careful when assuming that growth in output is representative of growth in welfare. Easterlin develops this theory in his 1974 paper, where he finds little to no correlation between GDP per capita and subjective well-being. He concludes, in agreement with Duesenberry and Abramovitz, that to judge their happiness, "people tend to compare their actual situation to a reference standard or norm, derived from prior and on-going social experience" (Easterlin, 1974).

In his 1995 paper, Easterlin goes on to find support for the second claim of the relative income hypothesis, by investigating this change within a country over time and sees similar results; that this reference standard is also based upon previous individual experience. Looking at more current literature, this supposition is supported by Ferrer-i-Carbonell (2005) who, looking at individual German panel data, finds that the income of the individual's reference group was just about as important for happiness as was the individual's own income. This position is also supported by McBride (2001) who conducts his study using American cross-sectional data. He claims that this effect may be lessened in low-income households. A slightly different theory relates to income aspirations; Stutzer (2004) finds evidence that higher income aspirations have a negative impact on utility. This aspiration point is again based upon individual experience and social comparison.

If we look at the evidence for this on a country-level, many papers find no correlation between GDP per capita and average levels of happiness, both between and within countries over time, as is apparent in Easterlin's 1974 and 1995 papers. This lack of relationship has been recreated since then with different datasets (Layard, 2005; Clark, Frijters & Shields, 2008). Kenny (1999) finds a weakly positive correlation in this relationship. However, upon including further control variables, he finds that this relationship is more likely due to factors that co-vary with GDP per capita.

Veenhoven (1991) criticises these claims, stating that Easterlin's results were incorrectly specified, and his dataset was not large enough. When he recreates Easterlin's cross-country (lack of) relationship, Veenhoven uses a more recent and larger world survey and with this data, he finds a correlation between GNP per capita and an average happiness of 0.84. As for Easterlin's positive relationship between

country income, Veenhoven also questions this, claiming that Easterlin's data was out-dated and claims that Easterlin significantly ignores the heterogeneities across countries. Again, he uses a more recent, larger dataset finding contradictory results. He finds that this positive correlation is only present in roughly half of countries, and the other half display only a weak and sometimes even a negative relationship, created by the economic prosperity of the country; a variable he claims is far more important to well-being levels. From these results, Veenhoven comes to the conclusion that, "people cannot be happy in chronic hunger, danger and isolation: not even if they have never known better and if their neighbours are worse off". This hypothesis is also supported by multiple other studies (Hagerty & Veenhoven, 2003; Di Tella & MacCulloch, 2008; Stevenson & Wolfers, 2008; Diener et al., 1993).

One of the problems with the paradox is the use of GDP itself. Many economists believe that GDP and similar indicators are overly relied upon and shouldn't be used as an indicator for well-being, as they don't allow for inequality, social inclusion or other important factors that contribute to happiness. The exclusion of these fundamental variables makes these economic performance indicators insufficient to measure well-being. Even Simon Kuznets, the developer of modern GDP, argued against the use of GDP as a measure of welfare or development. This view has been carried forward and in their paper 'Beyond GDP: The Need for New Measures of Progress', Costanza et al. (2009) sum it up quite nicely saying that, "useful measures of progress and well-being must be measures of the degree to which society's goals (i.e., to sustainably provide basic human needs for food, shelter, freedom, participation, etc.) are met, rather than measures of the mere volume of marketed economic activity". While we won't physically test for this within this study, it is something we will bear in mind and discuss in later sections.

A further problem we experience is that there is an assumption that all individuals experience a similar reaction to increased levels of GDP per capita; we must also consider that this effect may not be the same for everyone. Kahneman & Deatonthe (2010) consider the variation in emotional well-being has on the impact of income on happiness. They find that while emotional well-being rises with income, there is no more improvement after an individual receives an annual income of $75,000. From

this, they conclude that a low income will exacerbate poor emotional well-being that stems from negative events, such as divorce or loneliness. Mohanty (2014) looks at a similar type of problem. He considers the relationships between positive attitudes, income and happiness for young- and matured-adults in America. He finds that since a positive attitude has a larger impact than income, educational policy aimed at increasing happiness through income later in life should also aim to supplement traditional schooling with education in behavioural skills, and by promoting both, should increase happiness later in life. These two cases provide evidence that higher household income in itself cannot create higher life satisfaction.

This difference in impact can be seen across several factors. One of these may be cross-country fixed effects (such as social norms, culture, etc.) that will, in turn, have an impact on individual well-being. This phenomenon is investigated by Yul Lee, et al. (2000), who look at the difference in subjective well-being between Korean and Canadian students. These two countries are chosen because of their distinctly different cultures: Korea tends to be associated with a more collectivistic culture, whereas Canada a more individualistic one. This position is supported by a study by Diener & Fujita (1995) who find that an individual's happiness function is often different depending on their goals, because goals are inherently different across cultures. Yul Lee et al. find that there are differences in absolute levels of happiness between the two groups of students. However, they both unexpectedly use similar ordered set of criteria to measure their avowed happiness. This matches up with a previous study by Diener et al. (1995) who find that students in the Pacific Rim had lower average absolute happiness levels, although they also find that this is because they express different levels of (dis)-satisfaction in differing areas. For example, they find that income is a much higher predictor of happiness in the US. Lu, Gilmour & Kao (2010) look at the difference in values and happiness when comparing the Western (United Kingdom) and Eastern (Taiwan) cultures among students. In this case, they find the relationship between values and happiness is much stronger in Taiwan. Again, this is another area where there is disagreement on how income contributes to happiness.

If we also investigate this difference from a more financial standpoint, we can consider the differences across high- and low-income groups. Dynan & Ravina (2007) find positive impacts in both higher relative and real incomes, although a notable difference is that this effect is more pronounced for people with above-average incomes. This follows logically that those whose utility functions are more impacted by income are likely to be higher earners; therefore, we must question the direction of causality of these results. Until now, we have assumed that there is only one direction of causality; that a higher income results in higher life satisfaction. However, we cannot simply assume this. It is possible that on average a higher life satisfaction may lead to a more confident individual, which in turn is reflected in their job and wages. Easterlin touches upon this in his 1974 paper. He claims that the reasoning for subjective well-being is not convincing enough to have the impact on income that his data predicts and, therefore, it is unlikely to be a causal factor. Furthermore, economic concerns tend to be the foremost reason for personal well-being, suggesting directional causality from income to subjective well-being. Smith & Razell (1975) also find evidence for this belief, studying lottery pool winners and the impact it has on well-being, and using non-winners as a control group. They observe that on average they are happier than the control group, which suggests that there is likely causality from income to well-being.

Oswald et al. (2015) argue that, conversely, there is also evidence that suggests that happier people tend to be more productive. They present evidence that happier people tend to be up to 12% more productive than those who reported themselves as unhappy. Of course, this is not to say that productive people always earn more, however it is likely that higher productivity on average is rewarded in a pecuniary sense (whether through promotion, wage increase or bonuses). The authors admit that there is one large problem with their study; since this study was performed under laboratory conditions, it may not necessarily yield similar results in a business setting. This is not to discredit this theory, and indeed, there are other papers that attest to seeing this relationship. Edmans (2012) argues that job satisfaction can predict future stock price finding. He states that “companies listed in the “100 Best Companies to Work For in America” generated 2.3% to 3.8% higher stock returns per year” whilst controlling for reverse causality. Böckerman & Ilmakunnas (2012) use Finnish data to

show that an increase in job satisfaction in a plant by 1 standard deviation will, on average, increase value-added in manufacturing by 6.6%.

I would also like to briefly mention a popular psychological theory, "Set-Point Theory", also known as hedonic adaptation (Brickman and Campbell, 1971). This theory suggests that humans have an inherent set point of happiness and that movement away from that point is only temporary, since they will always move back towards it as they grow accustomed to their new standard. Brickman, Coates, and Janoff-Bulman (1978) conducted a qualitative study on "lottery winners and accident victims" and indeed found that often the happiness levels of these people would eventually return to their "normal" levels.

As shown in this section, there is still disagreement when it comes to this relationship and as such, there is no definitive answer to Easterlin's Paradox. This is largely due to supposed inconsistencies in his analysis. This paper shall assess one of these inconsistencies, and test whether Easterlin's use of GDP per capita was a sufficient enough proxy to make the large claims that he did. We shall test this by seeing whether other indicators, used as either substitutes or as complements to GDP per capita, are more appropriate. We shall, however, not be considering Easterlin's results on the individual level, and although many other papers do, it is not within the remit of this paper.

3. Data & Methodology

(i) Data

The dataset is taken from three sources. Firstly, the happiness data (as well as other controls, including health and education expenditure) come from the Gallup World Poll dataset. It contains 159 countries containing 1431 observations between 2006 and 2015. It is an unbalanced dataset as various countries were included and excluded in different years. Furthermore, different questions were only included later on and so various controls have a more limited timespan. The Gallup World Poll has commonly been used in similar studies to this one in the past, as they offer happiness data and is considered a highly reliable database. The database is built using either telephone surveys in countries where telephone coverage represents at least 80% of the

population, otherwise they use an area frame design for face-to-face interviewing in randomly selected households.[2] According to them, "the target population is the entire civilian, non-institutionalized population, age 15 and older. The coverage area is the entire country, including rural areas, and the sampling frame represents the entire non-institutional civilian population" (Gallup, 2014).

The question asked by the Gallup World Poll about an individual's happiness is:

"Please imagine a ladder with steps numbered from 0 at the bottom to 10 at the top. Suppose we say that the top of the ladder represents the best possible life for you, and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time, assuming that the higher the step the better you feel about your life, and the lower the step the worse you feel about it? Which step comes closest to the way you feel?"

This is important to note as the disagreement in the best way to assess subjective well-being/happiness/satisfaction can change depending on the question asked. Although we previously stated that we would be considering all of these phrases to be synonymous in previous literature, from now on, this study will be exclusively considering life satisfaction as the question asked in the Gallup poll seems to point towards this measure.

The second source is the World Bank database, which most importantly for us, provide different economic performance indicators that we shall be using, including Gross Domestic Product (GDP) per capita, Gross National Income (GNI) per capita, Net National Income (NNI) per capita and household consumption (HHC) as well as several control variables (inflation rate, unemployment rate, life expectancy, and CO2 emissions per capita). This dataset is again a country-level panel dataset, containing 220 countries (and 44 other aggregated regional data) containing 2640 observations spanning between 2006 and 2015. This is a weakly balanced dataset, as there are multiple missing observations for different countries in different variables. Although there is scope for this paper to use data pre-dating 2006 with the World Bank, we

[2] These countries reside mostly in Latin America, former Soviet Union countries, nearly all of Asia, the Middle East, and Africa.

shall not move outwith the dates of the dataset allowed by the Gallup database. Furthermore, the timespan of the dataset contains sufficient variance for our analysis.

The final dataset that we use is UN data, for its HDI indicator. HDI is a UN measure and therefore is only available from the UN itself. This measure will likely return the most dissimilar results as it is a composite statistic that is made up of life expectancy, education, and per capita income indicators. It gives a score from 1 to 10 and is described as, "a summary measure of average achievement in key dimensions of human development [in order to] emphasize that people and their capabilities should be the ultimate criteria for assessing the development of a country, not economic growth alone" (UNDP, 2016).

The following economic indicators have been chosen as each one can be used to predict the economic performance of a country on an individual level: GDP per capita, GNI per capita, NNI per capita, Household final consumption (HHC), and the UN's Human Development Index (HDI). Although these will co-vary in a similar method due to the high correlation between them all, there is still a level of variation between them all that will allow us to see differences in results (Boarini, Johansson & d'Ercole, 2006). Notably, HDI is formed in a completely different way to our other indicators and we are most likely to see the biggest variation in its impact. As each indicator varies distinctly, this is hugely useful in our search to see how economic performance impacts life satisfaction levels in a country; interpreting these indicators alongside each other can be complementary in the analysis and in fact, gives us a better picture of the relationship. How each indicator is built and how it varies from the others are provided in table I below.

We are investigating these specific indicators not for arbitrary reasons (although the availability of data, of course, has to be taken into account). GNI per capita and NNI per capita are both chosen specifically as they are highly comparable to GDP per capita. Evaluating the differences with these measures firstly can act as a check for the robustness of GDP, as they are likely to move in similar fashion and direction to the former. Secondly, since they are not analogous in every instance, we can still observe the effects that the above mentioned differences have on life satisfaction.

Theoretically speaking, there are several different impacts this could have. One possible difference would be if we also consider that GNI and NNI per capita take into account the freedom of each country to trade on the global market then as Gwartney & Lawson (2006) suggests. Therefore, the more open the country is to international trade, the more satisfied the citizens will be.[3]

Indicator	How it is built	Major differences
GDP per capita	GDP is the sum of gross value added by all resident producers in the economy, plus any product taxes and minus any subsidies not included in the value of the products divided by midyear population.	GDP is typically the most commonly used indicator of economic performance. It's usage, however, is so common, it can lead to use of it when it may be more appropriate to employ another indicator.
GNI per capita	GNI is the sum of value added by all resident producers, plus any product taxes (less subsidies) not included in the valuation of output, plus net receipts of primary income (compensation of employees and property income) from abroad divided by midyear population.	The largest difference between GNI and GDP is the "terms-of-trade" effects. GNI includes the value produced from all citizens, therefore this is adjusted for a country's performance on the global market as well as its internal development.
NNI per capita	Adjusted net national income is GNI minus consumption of fixed capital and natural resources depletion.	NNI accounts for the depreciation of capital stock, therefore it can be considered a more "real" measure.
HHC	Household final consumption expenditure is the market value of all goods and services, including durable products (such as cars, washing machines, and home computers), purchased by households.	This is not a measure of national output, but simply a measure of household consumption as opposed to including corporate, public, or external transactions and income.
HDI	The Human Development Index (HDI) is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and having a decent standard of living. The HDI is the geometric mean of normalized indices for each of the three dimensions.	While the HDI does take into account economic performances, it also includes multiple other dimensions that are important for its measurement. This therefore accounts for other factors that can impact life satisfaction and so we should see a different relationship.

Table I: Indicators of economic performance
Sources: The World Bank, United Nations Development Programme

[3] In this case we would consider a lack of competitiveness as a barrier to trade entry.

We should expect household consumption to be highly correlated with former indicators, however it will be different for several reasons. In an accounting sense, the main differences in this case, as opposed to our previous indicators, is in the measurement. Firstly, here we are only looking at consumption and ignoring all types of personal savings or business retained earnings (that aren't spread around the economy). This can have interesting results as different countries have rates of savings that can be attributed to both current micro- and macro-economic trends, as well as cultural and risk-level propensities which all impact the consumption and savings rates across countries (Edwards, 1995; Hofstede & Bond, 1988). Secondly, its measurement: whilst the others are measured on a country level before being adjusted by population size, household data is, as is suggested in the name, measured on a household level using representative country samples. This can have both positive and negative impacts on its comparable accuracy; unfortunately it may, in particular, lead to problems if the study sample is not representative of the whole population or if the respondents are subject to misreporting deliberately (due to illegal activities) or accidentally (due to inherent bias and mistakes). While we acknowledge that this issue can occur, but still think it is a valid and relevant measure to include as it allows us to also consider the marginal propensity to consume and to save in our data.

Finally, HDI is by far the most different item, as it incorporates other variables that aren't considered in the other indicators. Including this indicator gives us the opportunity to consider whether these types of measures (such as Bhutan's GNH) can accurately be used as gauges of satisfaction, especially as we can present this alongside the measures that are currently used, such as GDP or GNI, and see how comparable they are. In this case, the indicator is adjusted for other variables that theoretically are in line with increasing life satisfaction and so we would expect a more linear relationship (UNDP, 2016).

Easterlin, by solely considering the basic correlation graphs, fails to control for a potential endogeneity problem. The inclusion of control variables enables us to avoid this problem and should present a more accurate depiction of the relationship. These control variables have been chosen through evidences in the current literature and table II below explains the reasoning behind each one.

Variable	Reasoning	Paper
Inflation	Individuals have to carry nominal money balances for transactions purposes, that are negatively impacted by inflation and so it has an impact on real income.	Di Tella & MacCulloch, 2008
Unemployment Rate	The costs of unemployment depend on the expected duration of unemployment spells, i.e. the likelihood of becoming unemployed.	Di Tella & MacCulloch, 2008
Life Expectancy	Considering the theory of compensating differentials, jobs that risk a shorter life span pay better, ergo a longer life expectancy is preferable on average.	Di Tella & MacCulloch, 2008
CO2 Emissions per Capita	Environmental degradation can have adverse effects on individual utility, as people assign a positive value to a cleaner environment.	Di Tella & MacCulloch, 2008
Education Expenditure	Everything else being equal, the marginal utility of additional income is higher for less educated people.	Castriota, 2006
Health Expenditure	People's expectations for health standards influence their reported health and associated life satisfaction, similar to the Easterlin paradox.	Graham, 2008.
Year Fixed Effects	Year specific phenomena can impact upon levels of life satisfaction and as such it is important to control for these outside of indicator changes.	-

Table II: Control Variables

(ii) Evidence

We have thus far assumed that life satisfaction is positively correlated with economic performance and as such, has a positive relationship with all of our selected indicators. Figure 1 below shows that this assumption is justified.

When we look at the relationship between our indicators and life satisfaction levels, we can see a clear picture in all cases. There is a strong positive correlation between better economic performance and higher self-reported life satisfaction across all of our indicators in all periods.[4] These correlation graphics give us good reason to investigate this relationship further as they directly contradict Easterlin's claims.

[4] See Appendix II for breakdown across time

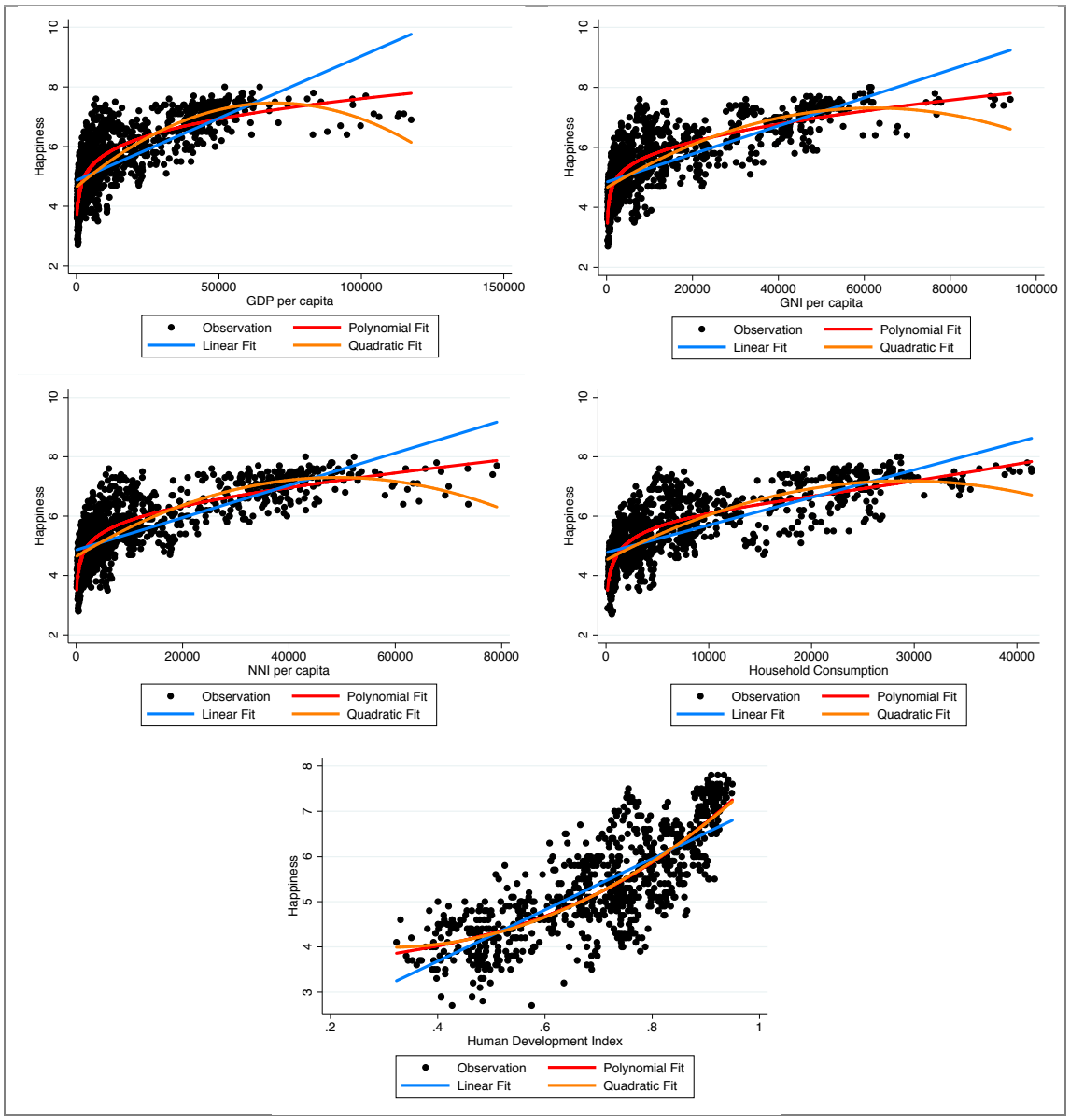

Figure 1: Happiness by economic performance split by economic indicator

Furthermore, we can see that largely these graphs show a similar fractional polynomial relationship over time, except in the case of HDI, where we see a more straightforward linear correlation (as predicted earlier). This isn't wholly surprising once we look at the correlation presented between our indicators. As seen in table III, these indicators are, in the main, closely correlated with one another. Table III shows the correlation between each of the indicators when GDP per capita = 1. Each of these relationships are largely similar, with GNI and NNI very close to 1. As we predicted previously, household consumption is slightly less correlated, considering that we have to take into account the differences caused by MPC and MPS. Finally HDI has a correlation of 0.7176; the lowest for previously discussed reasons.

Indicator	GDP	GNI	NNI	HHC	HDI
Total	1	0.9878	0.9957	0.9312	0.7176

Table III: Indicator Correlation

We can also compare these correlations over time and between regions (as determined by the World Bank). Completing these regional- and time-specific analyses allows us to check the robustness of our results as we may see varied regional- or time-specific effects that can distort our results. As shown in Figure 2, we see that these correlations are not necessarily the same across all regions. Unsurprisingly again, GNI and NNI vary the least, with household consumption again having slightly more variation and HDI the largest differences between regions. Notably, these variations are slightly different and more extreme in the North American region. However, it is highly likely that this difference in North America is exaggerated because it has so few observations (only 2 countries[5]). Whilst we observe some variation between regions, the variation between years, however, is unnoticeable. This suggests that although these indicators may not move exactly in line with one another, over time there is little change in how they vary.[6]

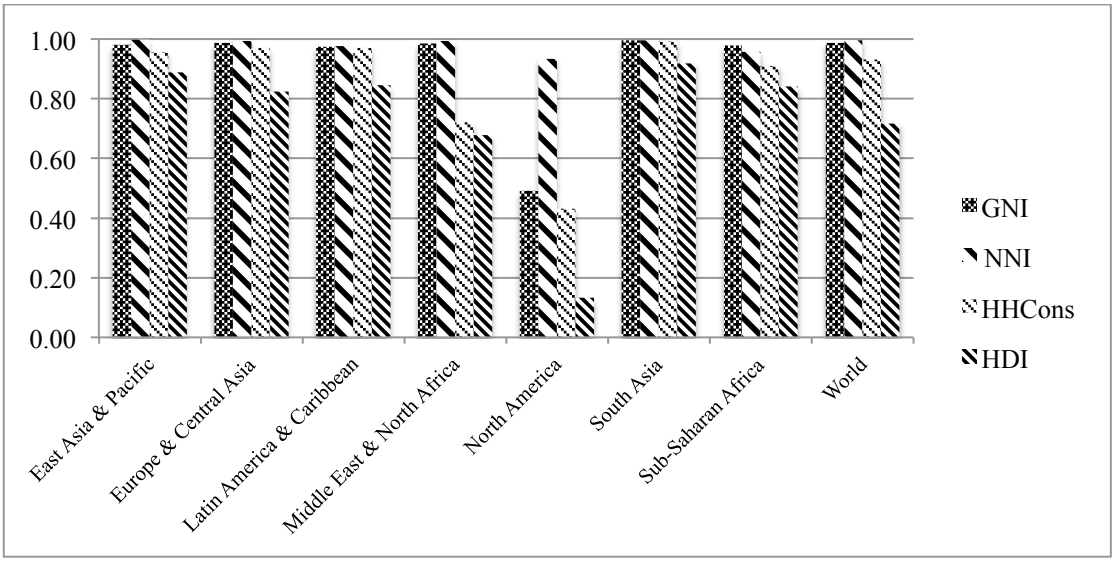

Figure 2: Correlation between indicators by region

[5] Canada and the United States of America

[6] We only go as far back as 2010 as we do not have data for HDI prior to that date, however the correlations for previous years are also similar for the other 3 indicators.

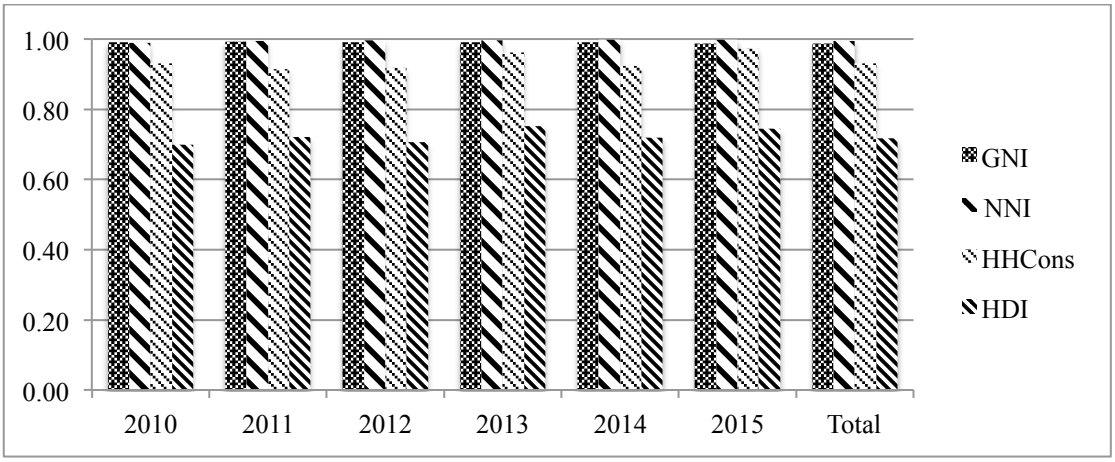

Figure 3: Correlation between indicators by year

We, therefore, look at the relationships between these indicators and life satisfaction between regions as well. Here, we see far more variation. As shown in Figure 4, we see huge differences between regions, but few differences between indicators.[7] For example, if we focus specifically on the East Asia & Pacific region, we see that the indicators move largely concurrently with one another. Meanwhile, if this is compared to the Middle East & North Africa region then, we can observe huge differences in each of these indicators over time between regions. However, again in this region the indicators move together.

Should we consider the actual correlations in Figure 4, it is now fairly obvious that they are dependent on the region. Our indicators, for instance, move alongside life satisfaction in East Asia & Pacific and Europe & Central Asia regions which suggest that these are good indicators of satisfaction, but, there is however, no clear correlation between the two curves in the Middle East & North African, North America, and South Asia regions. In Latin America & Caribbean and Sub-Saharan Africa regions, there are elements of correlation but it is not so clear-cut as others. These regional differences (as opposed to time differences) could be presented as evidence for cultural or social differences creating variation in the impacts of our indicators across regions that were previously hypothesised by the aforementioned Diener & Fujita (1995). Should these claims be true, this would be an area that was

[7] Due to the number of graphs required for figure 4, they have been reduced to not include information like indices or keys and they simply show the indicators (red) over time compared to the life satisfaction (blue) over time.

deficiently tested by Easterlin, and would indicate that his claims were not supported in all cases under sufficient evidence.

Although we have said that there are few differences within countries and between indicators, the one exception to this would be HDI. HDI shows an upward trend throughout all of these figures regardless of satisfaction levels. However, it is also very important to remember that the indices are not similar in these graphs as our aim is to capture the similarities in change over time with these indicators and life satisfaction and not to compare whether these absolute changes are absolutely comparable with one another. Whilst we cannot necessarily judge the gradient slopes against one another, we can say that despite an improving HDI in all regions over time, this is not necessarily reflected in improving levels of life satisfaction.

(iii) Hypotheses

From the evidence presented above, we shall propose the following hypotheses:

- **There is a positive correlation on average between GDP per capita and life satisfaction**

In order to test this hypothesis, we shall use a controlled regression model to test whether there is a significant positive coefficient when we regress GDP per capita (in $100,000) on life satisfaction levels across all available regions and time periods. Therefore the null and alternative hypotheses for this are:

$$H_0: \beta_{GDP\ per\ capita} > 0$$
$$H_1: \beta_{GDP\ per\ capita} \leq 0$$

- **This relationship shall be shown to be robust when using alternative indicators of life satisfaction**

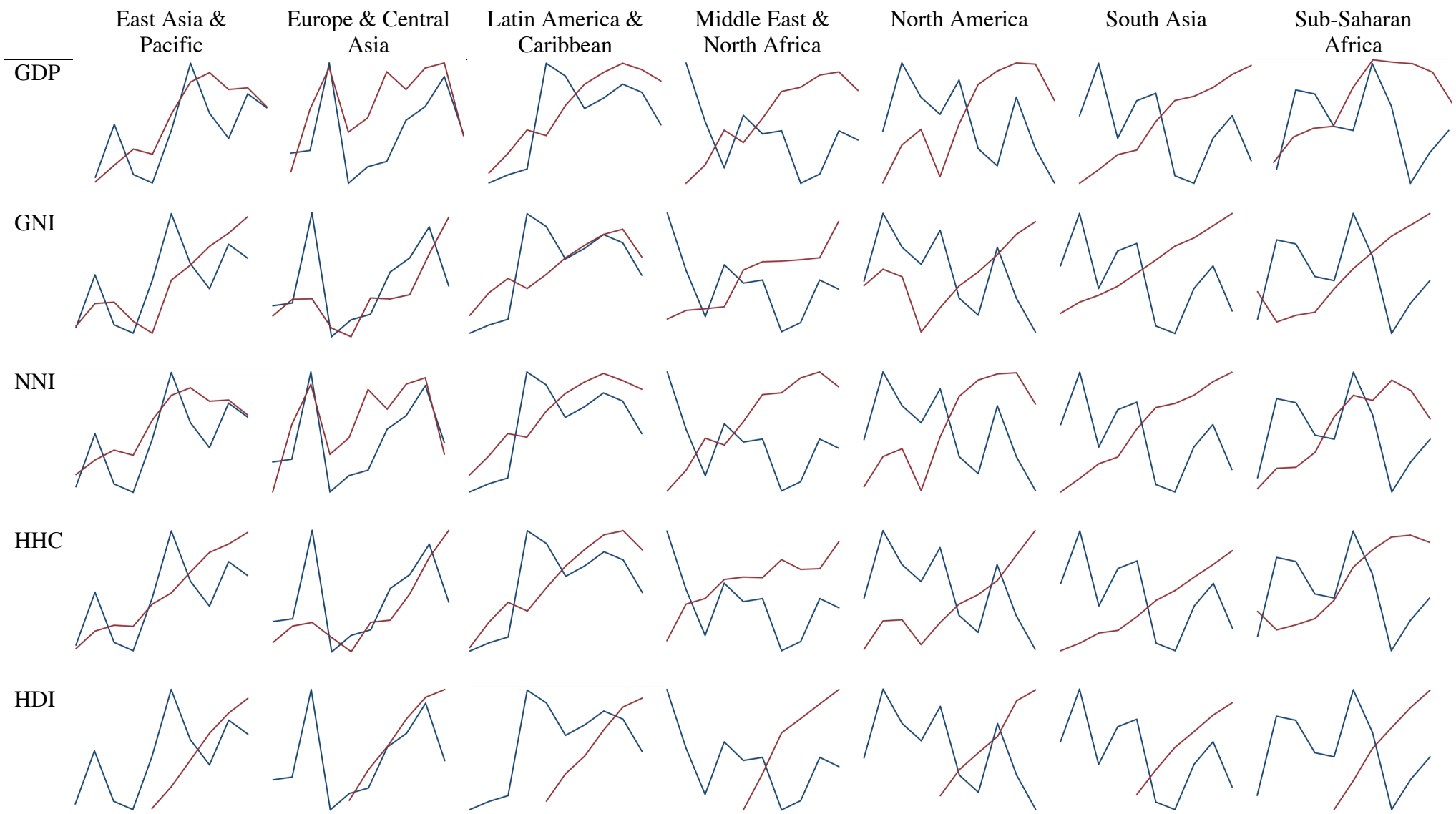

Figure 4: Correlations over time of indicators (red) and life satisfaction (blue)

For our second hypothesis, we shall be analysing multiple variables of interest and using the same model as with the initial hypothesis. However, we shall replace GDP per capita with each of the other respective indicators (GNI per capita, NNI per capita, Household Consumption, and HDI). In this case, we shall be testing whether the coefficients for each of these indicators is similar in direction as with GDP. Therefore, our null and alternative hypotheses here are:

$$H_0: \beta_{GDP} \approx \beta_{GNI} \approx \beta_{NNI} \approx \beta_{HHC} \approx \beta_{HDI}$$
$$H_1: \beta_{GDP} \not\approx \beta_{GNI} \not\approx \beta_{NNI} \not\approx \beta_{HHC} \not\approx \beta_{HDI}$$

- **This relationship shall be varied when comparing different regions**

This final hypothesis is similar to the previous in that we shall be comparing the coefficients between both indicators and now also regions. Therefore, the null and alternative hypotheses are:

$$H_0: \beta_{I1,R1} \not\approx \beta_{I1,R2} \not\approx \beta_{I2,R1} \not\approx \beta_{I21,R2}$$
$$H_1: \beta_{I1,R1} \approx \beta_{I1,R2} \approx \beta_{I2,R1} \approx \beta_{I21,R2}$$

where I1 and I2 are different nonspecific indicators and R1 and R2 are different nonspecific regions.

(iv) Statistical Analysis Methods

In order to analyse these relationships further, we use random effects models. We have chosen to use this type of model over a pooled OLS regression as the availability of panel data allows us to control for unobserved heterogeneity, so that the models cannot be biased by omitted time-invariant characteristics. We have also chosen to use it over a fixed effects estimator, as it would only allow us to analyse the within-heterogeneity, as opposed to being able to look at both within- and between-heterogeneity like we do in this analysis. The importance of being able to capture both within- and between-variation is integral to this study; the variation across countries and regions means that we are likely to experience a difference in effect and so using

a fixed effects model would nullify our ability to compare this.[8] In order to show the robustness of these results, we shall also complement our random effects models with pooled OLS and fixed effects models within our discussion section to compare the results. To further prove the robustness of the results, we shall also be creating these models, staggering the inclusion of controls to show the effect of including different controls has on the variable of interest. As each of our regressions also considers a different variable of interest, they will each be considered in separate segments in our results section.

4. Results

(i) Hypothesis 1

In order to test our initial hypothesis, we regress GDP per capita on life satisfaction levels before adding in further control variables to avoid an endogeneity problem. Table IV initially shows a positive relationship between GDP per capita and satisfaction, but when these control variables are introduced, this relationship becomes insignificant. This would suggest that GDP per capita does not has a positive correlation with life satisfaction; a direct evidence against our null hypothesis. However, a good reason for this is likely to be that one of our control variables is highly correlated with GDP per capita and as a result captures the effect that was initially shown in regression (1). Upon further analysis, we run regressions excluding individual controls one at a time to check whether one or more of them will make the effect of GDP significant again. We find that omitting Government Health Expenditure from the regression makes the effect significant and positive to the 1% level. From this we check the correlation between these variables and discover that the variables have a correlation of 0.9197. This would definitely be cause for such a change. Therefore, regression (4) shown below has omitted this variable in order to avoid a multicollinearity problem. Subsequently, we see a significantly positive effect, and as a result, fail to reject the null hypothesis.

If we look back at the correlation graphs, we also can identify that GDP per capita and life satisfaction do not have a linear correlation and that it seems to peak off at a point

[8] We use this random effects model despite running Hausmann tests (Appendix III) that suggest a mix of both. Although this may present slight bias in our results, this can be taken into account but we cannot observe between heterogeneity in fixed effects regressions, which is more important.

just under $70,000[9] if we look at the relationship in Figure 1. Therefore, in model (5) below we have included a quadratic element. Here we do, in fact, see a negative coefficient for the quadratic variable. This would suggest diminishing returns from GDP per capita. This is in line with multiple other theories that suggest there is a saturation point from income (Cummins, 2009). Whilst logically a saturation point makes sense (due to an upwards bounds of life satisfaction between 1-10), I am hesitant to sponsor this claim based on our results; in Figure 1, it is clear that our fractional polynomial curve is a better fit and has a continuously upwards sloping curve.

Table IV: Hypothesis 1

Variables	(1) (Only GDP)	(2) (Added Controls)	(3) (Full Model)	(4) (Excl. Health)	(5) (GDP-Squared)
GDP per capita ($100,000)	3.021***	0.503	0.707	1.821***	5.058***
GDP per capita^2	-	-	-	-	-3.433***
Unemployment Rate	-	-0.029***	-0.025***	-0.024***	-0.023***
Life Expectancy	-	0.056***	0.065***	0.064***	0.048***
CO2 emissions per capita	-	0.034**	0.021	0.013	0.002
Inflation Rate	-	0.000	-0.000	-0.000	0.001
Govt. Education Exp.	-	0.021***	0.023***	0.021***	0.021***
Govt. Health Exp.	-	0.000**	0.000***	-	-
Time Fixed Effects	No	No	Yes	Yes	Yes
Constant	4.987***	1.046*	0.461	0.557	1.469**
Observations	1,248	637	637	643	643
Number of Countries	158	126	126	127	127
Within R^2	0.0125	0.0403	0.0476	0.0505	0.0667
Between R^2	0.5464	0.7100	0.7366	0.7095	0.7370
Overall R^2	0.5093	0.7078	0.7267	0.6974	0.7158

*** p<0.01, ** p<0.05, * p<0.1

With regards to the control variables, there are few surprises as most directions of variables move as expected or are insignificant. However, as they are simply included

[9] (8.077003/11.581842) * $100,000 = $69,738

to use as controls for our variables of interest, we shall not be analysing them separately.

(ii) Hypothesis 2

To test our second hypothesis we have recreated model (5) from Table IV, using each of our alternative indicators (also continuing to omit health expenditure as it is highly correlated with each of our other indicators). If we look at the results a similar pattern emerges in most of them. There is a positive significant linear coefficient and a significant negative quadratic coefficient, except in the case of HDI where we see significant negative linear and positive quadratic coefficients. Once again, This is aligned with Figure 1, where this type of relationship between HDI and life satisfaction is noticeable. If we wish to look at an only linear model for each of these cases (similar to model (4) in Table IV available in Appendix I), then it is obvious that all of these indicators have significant positive relationship. Therefore, from Table V we partially fail to reject our second hypothesis. Whilst HDI breaks the mould and doesn't fit a similar curve as the others do, it does still have a constant positive correlation within our bounds that shows a level of robustness to our analysis.[10] Although the sizes of the coefficients are different in each model, this is because each indicator measures a slightly different subject; however, this is irrelevant to the fact that we see this pronounced positive (decreasing) relationship.

Another point worth mentioning here is the change in Adjusted R^2 with all of these indicators. We can see that each of our alternative measures is a "better" within fit than GDP per capita. However, in terms of between variation, only GNI, NNI and HHC are an improvement and past this, HHC and NNI are better overall fits. Of course, this does not tell us which indicator is best, but considering them as a measure of best fit is highly relevant when we judge which is best used to analyse them as a measure of life satisfaction. Therefore, this could suggest that some of these indicators would be more appropriate as a proxy than GDP per capita when it comes to looking at the link between economic performance and life satisfaction, i.e. HDI

[10] HDI may not have exactly similar coefficients to our other indicators, but when we look at the top and bottom bounds of our happiness levels (1-10) it has a consistently positive correlation, similar to all of the other indicators.

when looking within countries, HHC between countries and NNI per capita when looking between and within.

Table V: Hypothesis 2

Variables	(1) (GNI)	(2) (NNI)	(3) (HHC)	(4) (HDI)
GNI per capita ($100,000)	6.755***			
GNI per capita^2	-5.098***			
NNI per capita ($100,000)		6.754***		
NNI per capita^2		-6.091***		
Household Cons. ($100,000)			13.089***	
Household Cons.^{2}			-19.696***	
Human Development Index (1-10)				-8.234***
Human Development Index^2				10.520***
Unemployment Rate	-0.026***	-0.024***	-0.024***	-0.034***
Life Expectancy	0.039***	0.050***	0.036***	0.011
CO2 emissions per capita	-0.005	-0.001	0.007	-0.008
Inflation Rate	0.002	-0.000	0.002	-0.003
Govt. Education Exp.	0.020***	0.022***	0.023***	0.034***
Time Fixed Effects	Yes	Yes	Yes	Yes
Constant	1.948***	1.369**	2.024***	4.865***
Observations	616	622	617	348
Number of Countries	118	123	119	115
Within R^2	0.0851	0.0713	0.0850	0.0973
Between R^2	0.7430	0.7572	0.7608	0.7282
Overall R^2	0.7080	0.7354	0.7204	0.7093

*** p<0.01, ** p<0.05, * p<0.1

(iii) Hypothesis 3

The most important feature of Table VI is the variability across both indicators and regions. This gives support for our third hypothesis. Easterlin's 14 countries consisted of a fairly well spread out make-up of countries.[11] This variability may, in fact, act as an impediment to seeing accurate results. Table VI below shows that there are several regions in which GDP per capita is insignificant and although this is also conversely

[11] East Asia & Pacific 2; Europe & Central Asia 3; Latin America & Caribbean 4; Middle East & North Africa 2; North America 1; South Asia 1; and Sub-Saharan Africa 1

true, to ignore the difference between regions can lead us to spurious results by generalising all regions as being similar. Therefore, these regressions are integral to our analysis as it confirms our third hypothesis; that the relationship between these indicators and life satisfaction is contingent on the region or area in which the individual is. For example, one interesting feature is the direction of these coefficients. There is (almost) a complete accord in positive correlations, aside from South Asia, where each coefficient has a negative relationship. Although only one of them is significant at even the 10% level, this is still a curious result as we have continually seen a converse relationship in all other cases. Asian culture has been posited to be different in the literature, but what our analysis still does not tell us though is the mechanism employed within the relationship. Whether this impact is cultural, social, or economic is something that can be analysed in more detail in the future.

What was not expected is the variation of indicators within each region. Our previous graph analysis showed a much more conjoint relationship between the indicators (aside from HDI). One reason for this may be that once we begin to control for other effects, these may correlate with our indicators. This could also be true in a specific region. For example, consumption in Sub-Saharan Africa is more likely to be correlated with unemployment rate than in East Asia & Pacific region, as the difference in unemployment in these regions is between absolute and relative poverty and this will have huge disparities on consumption levels. This is important to see as it shows us that one indicator does not necessarily give us the best indicator in all cases, even a composite indicator like HDI that is built in order to measure development and economic performance.

So which indicator is the best to use? There is no correct answer to this question. All of our indicators have their respective merits and weaknesses. In this analysis, we have included all of them, not to try and distinguish which is preferred, but to provide a level of robustness that hasn't been provided in other studies. Additionally, including them all gives us a more rounded view of the picture; these results tell us what is more impactful in specific regions.

One significant issue within the analysis is the lack of observations after we have split them up into smaller subgroups by region, as is more evident in region 5 (North America) where we have been unable to produce results due to a lack of observations. The range of observations for these regressions is between 18-250 with 21 out of 30 regressions under 100. The low number of observations in each subgroup means that the results are likely to be more biased and we are less likely to see significant results.

Table VI: Indicator Coefficients in Controlled Linear Models per Region[12]

Measure	(1) (East Asia & Pacific)	(2) (Europe & Central Asia)	(3) (Latin America & Caribbean)	(4) (Middle East & North Africa)	(5) (North America)[13]	(6) (South Asia)	(7) (Sub-Saharan Africa)
GDP	1.283**	0 .853	8.360***	3.704**	-	-18.941	4.713
GNI	-0.050	2.685***	17.443***	7.231***	-	-12.407	8.024
NNI	1.301	1.860***	10.362***	2.939	-	-17.569	10.352
HHC	0 .122	4.895***	22.658***	25.979***	-	-7.541	22.094**
HDI	5.301	5.274**	11.923***	9.430	-	-12.534**	3.189**

*** p<0.01, ** p<0.05, * p<0.1

5. Discussion

(i) Discussion of results

We find strong evidence for two out of the three (hypotheses 1 and 3) that, on average, economic performance is positively correlated with levels of life satisfaction and that there is variation in this relationship between regions. With the other (hypothesis 2) we find weak evidence that economic performance indicators largely give us comparable if not exactly analogous results. What does this tell us about our initial research question: does this provide robust evidence for the relationship between economic performance and life satisfaction levels? Furthermore, what impact does our result have on the Easterlin Paradox?

Our results suggest that on average there is a positive relationship between levels of life satisfaction and the economic performance in a country. This is exactly as many would expect. However, it is not necessarily as simple as we would expect as we also see much variation between indicators and between regions. This variation, shown in

[12] Total regressions in Appendix IV.

[13] North America has been omitted from our results due to insufficient observations.

Table VI, proves that different populations consider different variables to be more important to their life satisfaction. Therefore, whilst on average there is a case for this positive relationship, there is so much more that needs to be taken into account in each individual case. This is an important result as it neither agrees nor disagrees with previous studies, but also explains why they may simultaneously find opposing results, depending on their dataset.

If we apply our results to Easterlin's analysis, we do not find similar results and we also see potential reasons as to why Easterlin may have found the results he did. Firstly, if he overestimated countries in which there is not a positive impact, then this would skew his results, unfortunately misrepresenting the "world average". Secondly, while there is no problem with his use of GDP per capita, his decision to check the robustness of this result means that, in my opinion, his results are left in doubt. This is not to say that our results are perfect as we still experience a plethora of problems. However our next two sections will discuss these problems as well as including a more formal robustness checks section in order to confirm that our results are indeed robust.

(ii) Use of GDP

As previously discussed, the use of GDP as a measure is obviously hugely important to our study and it is worth discussing why GDP may not be a good measure for life satisfaction, happiness or well-being, what it's use means for the study, and also potential alternatives.

The intuition behind using GDP as a measure of well-being is that since GDP per capita is a gauge of a country's capacity to deal with the needs of its residents, increases in GDP per capita can be used to meet higher levels of the population's needs, and thus for their greater well-being (Ladaique, 2007). However, as was mentioned in our literature review, several high profile economists have questioned the legitimacy of using GDP or GDP per capita as a measure of welfare. This is largely due to fact that while the monetary value of all goods and services produced in a country is probably highly correlated with the welfare of a country, it should not be treated as synonymous. It is likely to be highly correlated for two main reasons:

firstly, a higher GDP per capita will have a causal impact on creating a happier population (for example, it is more likely that individuals have a comfortable lifestyle and have a better quality of life, both of which are associated with a higher income); secondly, it will also have a non-causal correlation, as a high GDP per capita is likely to be correlated with other variables that will result in a better level of welfare such as better institutions or quality of available education.

However, despite this high correlation, it is still difficult to argue that it is a good measure of welfare. This is because so many other factors should also be included in such a measure. The welfare and happiness levels within a country are highly contingent on various other factors that are not counted within the GDP measure. For example, two factors that are often cited are income equality and social cohesion. Whilst GDP per capita will tell us how a country's citizens compare on average between countries, however this does not tell us anything about the spread of the wealth. This is incredibly important: if, hypothetically, one individual owns all the wealth within a country, then this will likely have a huge effect upon happiness levels amongst its citizens. One way to compare countries on this level would be by using the Gini coefficient. However, this would again have to be used complementarily, alongside other factors, because if we were to use the Gini coefficient only, we would face the same problem, as before whereby we only account for one factor (in this case income equality) and ignore others.

Social cohesion is something that is completely ignored within most economic indicators. However, it is one of the most important factors to life satisfaction. Ladaique (2007) tests multiple social indicators to see their similarity to GDP. He looks at several different indicators, including social cohesion. He uses four different variables to determine the social cohesion of a country: volunteering, victimisation rate, prisoners, and suicide rates. He finds that while there is a positive correlation with volunteering work and GDP, there is no significant correlation with the other three. This shows that there are some levels of social cohesion that are not captured in GDP per capita; this is evidence that complementing economic performance indicators with social indicators provides a fuller picture.

One further problem with the use of GDP is the fact that certain events can be misrepresented with regards to welfare within a country. Firstly, unpaid volunteer work or casual favours between friends (i.e. off the books work), which create value in an economy, is not included in the measure. It can even lower the value of GDP, as it replaces the opportunity for paid work for another individual. This is obviously difficult for us to adjust for, as there is rarely a record of these transactions. This is also an issue when measuring welfare since these actions are often intimate or personal and can have impacts on life satisfaction. A second misrepresentation is that GDP per capita can actually be hugely influenced by negative events, such as war or natural disasters. This works in two ways as GDP is measured per capita and so, the significant reduction of a population will increase the 'mean' income. Furthermore, events such as war require a huge investment in the economy, for example, the GDP of Germany increased substantially in the lead up to World War 2. However, whilst these types of events can lead to an increase in GDP per capita, they also often lead to a significant decrease in the welfare of the population and levels of life satisfaction are likely to fall dramatically (Veenhoven, 1991).

One of the main issues for these alternative indicators, however, is also one of the main benefits of using GDP per capita; the availability of GDP data is almost guaranteed in most cases, whereas finding data on alternative indicators can be very difficult. It is also incredibly complex for individuals or companies to be able to gather this data first hand, due to the breadth of questioning that it would involve and the timescale that would be required. Therefore, in order to complete studies such as this one, GDP per capita is a very useful tool despite its downsides.

This is all of course highly relevant to our study, as using our alternative indicators was intended to show that analysing different indicators alongside would complement the use of GDP. Although our indicators were largely similar to GDP, we did show that they could be used to highlight key areas of variation in different regions. Using even more dissimilar indicators may give us a more accurate view of the welfare of a country. Therefore, it is worth using multiple indicators simultaneously to be able to fully assess the satisfaction levels of a country, as simply using one gives an insufficient resolution. This is also a key factor in the 'Easterlin Paradox'. Due to the

sole use of GDP per capita in Easterlin's initial analyses, he was unable to provide a more comprehensive picture of well-being levels in a country and consequently his results were biased and fallible. However, if he had complemented GDP with other factors, such as the Gini coefficient and social inclusion amongst others, he may have seen different results.

(iii) Robustness Checks[14]

The first two models we shall take a look at are pooled OLS models and fixed effects models, found in Appendices V and VI. We analyse these regressions in order to see if the results are altered by the different circumstances that these models present. With a pooled OLS model, we would not expect to see much difference as we still observe both between and within variation; however, the main difference here is that in using panel data we are able to exploit the correlation between the error terms, and combine both the within and between variation efficiently.

The results are largely as we would expect: the pooled OLS models are almost identical in terms of direction and significance to our random effects model. However, the fixed effects model is very different. There are two main differences within the fixed effects models; firstly, GDP per capita, NNI per capita and household consumption models do not show significant results, and then perhaps more surprisingly, there are no significant results when looking at our third hypothesis regressions. If we initially look at the difference in significance for our first two hypotheses, we must consider what this result means. It indicates that within a country only a change in GNI or HDI will have a significant impact on life satisfaction levels. This is likely due to small differences in the makeup of these measures that impact internal features within a country, such as trade flows in GNI or education in HDI. For the secondary change, while this also reflects this lesser change within a country, we also experience a huge problem with the number of observations in these regressions. The lack of variance within these countries by region makes it very difficult for us to draw any significant results. Therefore, although we must acknowledge that these fixed effects models show an important difference when we

[14] All of the regression tables within this section will be included in the Appendix and shall be clearly indicated which Appendix throughout.

solely consider within variation, these checks are not significant evidence against our initial results.

Although cultural differences by regions are something that we have investigated, another important way to split our dataset is by income levels. There have been several studies that claim low-income households are much more responsive to increases in income, so that there is a diminishing marginal return in happiness from income. The World Bank data also provides a categorical variable that indicates whether each country is considered to be low-income, middle-low-income, middle-high-income, or high-income. Therefore, we have run our model split by each category using GDP per capita as our economic performance indicator, and the results are displayed in Appendix VII.[15] This test was not included within our main results, as it is not something that is not within the remit of this paper. However, it is included within this section as it could explain whether there are any regions that are affected by differing income levels instead of cultural differences. However, the results are fairly surprising; we see a converse relationship where all relationships are significant except for the low-income group.[16] Initially, this suggests that the average life satisfaction levels are less likely to be impacted by an increase in economic performance. Again, however, this may simply be correlatory. The low-income subgroup almost exclusively consists of countries in Sub-Saharan Africa, and as such it is difficult to establish whether it is because they are low income or from Sub-Saharan Africa or even another factor that causes this effect.

In order to run complete robustness checks, we must also consider physical checks. One problem that we must consider is the presence of stationarity in our happiness dataset. In this case, we would expect stationarity if we believe there is a reason for a linear trend within our dataset. We can check for this by using a Dickey-Fuller test; this test checks whether a time series variable is non-stationary and possesses a unit root, which would then suggest that happiness follows some kind of path and therefore that our results are biased. The results from this test are in Appendix VIII. They suggest that we should fail to reject the null hypothesis that there is a unit root;

[15] This model is run as a pooled OLS model, simply for simplicity.

[16] Although only GDP per capita is shown in the appendix, all other indicators displayed similar results aside from HDI where neither the low-income nor the middle-low-income groups were significant.

therefore, our happiness data may suffer from a lack of stationarity. However, if we boil this down further we are able to check this on a regional level, where we fail to reject at the 10% level for only 4 out of the 7 regions: East Asia & Pacific, Latin America & Caribbean, North America, and Sub-Saharan Africa. Whilst the presence of non-stationarity is not ideal within a study, it is not necessarily unexpected and does not take away from our results. Furthermore, we know that this is not present in all regions and therefore, can conclude that it is not a problem within our data, but may just be due to other reasons. It could potentially be a result of a natural occurrence in the dataset, such as the alleviation in 2008 credit crunch creating a steady upward happiness trend in North America, or due to improving health and education conditions in Sub-Saharan Africa. This is even more possible as we only consider a 10-year timeline and so, we do not know if this is a long-term trend or just something that has occurred within this short period.

One final check we shall carry out is the difference we see in results when we interact a categorical region variable with our indicator variables, as opposed to splitting the dataset into seven separate regressions (Appendix IX).[17] The results here are mostly as we would expect, they tell us that each of our indicators remain significant and keep the same directions. On top of this, we also see that each indicator by region still acts differently as we saw in our initial results. Although some of the significance for several of the coefficients is different, this is not hugely surprising as they are all now compared to a base level, differently to when we split the dataset.

(iv) Limitations

One limitation that can occur with any quantitative study is regarding the quality of the data. This can be even more important when we look at happiness data, as it has to be completed using self-reported data. The biggest fundamental issue with this type of data is that while people understand the concept of happiness, it is a difficult idea to model or give a cardinal numeric value to (White, 2014). As a result, we see huge disagreement in the best way in which to question people over their subjective happiness or well-being, as using an inadequate question or survey can lead to biased results. Bertrand & Mullainathan (2001) claim that there are three main problems

[17] These models are also done using a pooled OLS model for simplicity.

from using self-reported variables: 'cognitive biases' can occur when people react to the information presented in front of them, and so can be easily manipulated by the line of questioning. For example, they may respond differently when a question is phrased positively or negatively; 'social desirability' is the want for respondents to appear positively in front of the interviewers, one famous example of this is in electoral polls when more people claim to vote than in the actual turnout; and 'non-attitudes, wrong attitudes and soft attitudes' are problems that occur when respondents do not feel strongly about a question or may even not understand their true feelings towards an opinion and as such may respond offhandedly, creating false information. This can generate large biases, especially in small samples. However, the Gallup World Poll, where our happiness data comes from, is a highly respected and reliable database and takes multiple measures to try and avoid the above-mentioned problems. Their methodology can be found on their website.[18]

A further problem that we continually face in our analysis is the availability of data. This can be seen in the control variables we include. For example, when trying to include health expenditure, we found that this is highly correlated with our indicators, but a separate variable that measured the performance of the health industry in an economy that was not so highly correlated would have been sufficient, had the data been available. As a result, we are unable to include this, along with multiple other controls that would have been interesting to include, should the data have been more available. This problem is not significant as the robustness of our results has been checked multiple times in different ways and so the inclusion of further control variables is unlikely to change our main results significantly.

Obviously, this lack of data points creates is that it severely reduces the number of observations within our regression and the efficiency of the models along with it. This could be seen as a major problem as data tends to be more readily available within higher performing economies, meaning that our omitted observations may not be randomly distributed. This may also be a problem that we experience in our robustness section when we split our dataset by income groups as we see far fewer low-income observations than in all other income groups.

[18] https://web.archive.org/web/20140902045224/http://www.gallup.com/strategicconsulting/156923/worldwide-research-methodology.aspx

One of the main reasons for conducting this analysis was to create Easterlin's analysis with a more "up-to-date" dataset. Although this was achieved, it would have been preferable to be able to conduct this analysis over a longer time period. However, again, due to data availability this was not feasible. Doing this over more years would allow us to complete a more efficient analysis. Looking at our 10-year analysis we do not see a large amount of change; a longer timespan may allow us to see more variation and if there are any significant differences from populations in different generations.

One final issue worth mentioning is the direction of causality between happiness and HHC. Whilst we previously discussed the direction between happiness and income, the causality between these variables is up for debate. It is difficult to argue that there are no increases in happiness as consumption increases since consumption is often caused by an attempt to increase happiness. Conversely you may argue that happier people are more sociable and have a higher propensity to consume, meaning that this relationship is incredibly hard to distinguish (Veenhoven, 2003). There are many papers that contribute to this disagreement but in this case it may be more prudent to consider simply the correlation between happiness and HHC, as opposed to the impact that HHC has on happiness. This is not necessarily a problem as we can still see how the two variables are correlated. However, the real problem occurs when this relationship is considered for things such as policy, since we cannot say with certainty that increasing consumption will raise life satisfaction levels in an economy or vice-versa.

(v) Policy Implications

In terms of government policy, this result can have multiple implications depending on the desired results. If we assume that the happiness of a country's population is the main aim of the government, then our results suggest that improving economic performance shall on average improve the levels of happiness in the country. An important feature to take from this is that the indicator used is key. As shown, HDI did have dissimilar results and because of this the composition of the indicator used would have large effects on the policy that should be pushed. For example, whilst

GNH may be a good composite indicator for Bhutan to use, other countries should consider what they are trying to achieve and then create a different composition that may be more appropriate for them.

This is even more pronounced when we consider the results from our third hypothesis. Across countries, we see large variation in reaction to different indicators. This suggests that countries should also look internally at what's important for their own country before thoughtlessly assuming that their policy should follow other countries'. For example, the South Asian region shows signs that economic performance is not correlated or potentially even negatively correlated with happiness; as a result policy in this region should consider focussing their efforts into other results aside from economic performance.

(vi) Further Research

Further research should focus on the mechanism used in the relationship across regions. The current literature largely suggests that these differences are because of different cultural attitudes towards economic performance. In order to test this, it would be best to run regression analysis to test how cultural differences impact the economic performance of a country. This can be a difficult to complete, as including cultural differences can be problematic in a regression analysis. Two potential solutions would be instead to analyse attitudes towards different cultural characteristics such as religion, social norms, arts, institutions, etc. A second line of thought would be to use an instrumental variable regression, using historical instants and data that predict how the current culture is by country, such as the one used by Acemoglu et al. (2005).[19]

As discussed previously, the availability of a wider data would have provided a more rounded analysis. This would be especially useful in recreating the regional analysis, where the small number of observations in some of the regressions means that it is far less likely that we see significant results. Therefore, it would be interesting to recreate this analysis with a longer dataset and include balanced control variables to increase the number of observations.

[19] Acemoglu et al. (2005) use settler mortality to predict current institutions in developing countries.

Finally, as I mentioned already, I would also recommend using multiple variables complementarily. This analysis only includes economic indicators, but for a more thorough analysis there are others that would also be interesting to consider. Ladaique (2007), for instance, considers multiple social indicators finding that levels of these indicators are highly correlated to GDP even if the changes are not. It would be an interesting research to consider these social, as opposed to economic, indicators to examine whether we still have a similar relationship to life satisfaction on average. Another measure to consider is GNH; as a composite measure specifically aimed to analyse happiness, it would be interesting to see how effective is it at determining the level of happiness. This will be especially important when running the analysis across regions, and taking into account the different cultural and social norms across these regions. Finally, median household income is a recent addition to the Gallup World Poll dataset, and it would have undoubtedly provided added depth to our analysis. Unfortunately, the timescale in the dataset did not correspond to the requirements of this paper.

6. Concluding remarks

In Easterlin's conclusions, he states that the only sure inference he takes from his study is that, "we need more research on the nature and causes of human welfare". The discussion created from his initial paper has worked to do exactly this and although there has been plenty more written on this topic, there is still more to explore. This paper aimed to consider Easterlin's own claims that there was no correlation between levels of self-reported well-being and GDP per capita. Having reconsidered this relationship, we have found evidence contradicting this claim and further indication that life satisfaction is on average correlated with economic performance in general and that this relationship is highly dependent on the region in which the country resides.

This result is especially important for a few reasons. Firstly, it explains why we see varied results in different studies as it tells us that the results are highly dependent on the examined population. Whilst this may seem like a banal conclusion, it is something that had been omitted from previous studies and is an important fact to

prove. Secondly, it shows us that GDP per capita and other individual economic performance indicators by themselves are not sufficient as gauges for well-being in a country. What may be more appropriate is to use multiple indicators, potentially also including social indicators or composite indexes to give us a more representative and detailed analysis.

To conclude, what this study explicitly tells us is that when we alter the circumstances around Easterlin's study of between country happiness and economic performance we find very different results. It is not evidence against or for the relative income hypothesis, but it does point out large flaws in some of the most important work for this theory.

7. References

Abramovitz, M. (1959). The allocation of economic resources (Vol. 17). Stanford University Press.

Adler, A. (2009). Gross national happiness in Bhutan: A living example of an alternative approach to progress.

Alkire, S., Ura, K., Zangmo, T., & Wangdi, K. (2012) A short guide to gross national happiness index. Thimphu: Centre for Bhutan Studies.

Argyle, M. (2013). The psychology of happiness. Routledge.

Ball, R., & Chernova, K. (2008). Absolute income, relative income, and happiness. Social Indicators Research, 88(3), 497-529.

Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data.

Bjørnskov, C. (2003). The happy few: Cross–country evidence on social capital and life satisfaction. Kyklos, 56(1), 3-16.

Boarini, R., Johansson, Å., & d'Ercole, M. M. (2006). Alternative measures of well-being.

Böckerman, P., & Ilmakunnas, P. (2012). The job satisfaction-productivity nexus: A study using matched survey and register data. Industrial & Labor Relations Review, 65(2), 244-262.

Brickman, P., & Campbell, D. T. (1971). Hedonic relativism and planning the good society. Adaptation-level theory, 287-305.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative?. Journal of personality and social psychology, 36(8), 917.

Castriota, S. (2006). Education and happiness: A further explanation to the Easterlin Paradox. Unpublished paper.

Costanza, R., Hart, M., Talberth, J., & Posner, S. (2009). Beyond GDP: The need for new measures of progress. The pardee papers.

Cummins, R. A. (2010). Subjective wellbeing, homeostatically protected mood and depression: A synthesis. Journal of Happiness Studies, 11(1), 1-17.

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. Journal of Economic literature, 46(1), 95-144.

Di Tella, R., & MacCulloch, R. (2008). Gross national happiness as an answer to the Easterlin Paradox?. Journal of Development Economics, 86(1), 22-42.

Diener, E., & Fujita, F. (1995). Resources, personal strivings, and subjective well-being: a nomothetic and idiographic approach. Journal of personality and social psychology, 68(5), 926.

Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. (1993). The relationship between income and subjective well-being: Relative or absolute?. Social indicators research, 28(3), 195-223.

Diener, E., Suh, E. M., Smith, H., & Shao, L. (1995). National differences in reported subjective well-being: Why do they occur?. Social Indicators Research, 34(1), 7-32.

Duesenberry, J. S. (1949). Income, saving, and the theory of consumer behavior.

Dynan, K. E., & Ravina, E. (2007). Increasing income inequality, external habits, and self-reported happiness. The American Economic Review, 97(2), 226-231.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. Nations and households in economic growth, 89, 89-125.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all?. Journal of Economic Behavior & Organization, 27(1), 35-47.

Easterlin, R. A. (2006). Life cycle happiness and its sources: Intersections of psychology, economics, and demography. Journal of Economic Psychology, 27(4), 463-482.

Easterlin, R. A. (2010). Well‐Being, Front and Center: A Note on the Sarkozy Report. Population and Development Review, 36(1), 119-124.

Edmans, A. (2012). The link between job satisfaction and firm value, with implications for corporate social responsibility. The Academy of Management Perspectives, 26(4), 1-19.

Edwards, S. (1995). Why are saving rates so different across countries?: An international comparative analysis (No. w5097). National Bureau of Economic Research.

Ferrer-i-Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect. Journal of Public Economics, 89(5), 997-1019.

Gallup World Poll. (2014). Worldwide Research Methodology. Retrieved May 26, 2017, from https://web.archive.org/web/20140902045224/http://www.gallup.com/strategicconsulting/156923/worldwide-research-methodology.aspx

Graham, C. (2008). Happiness and health: Lessons—and questions—for public policy. Health affairs, 27(1), 72-87.

Gwartney, J., & Lawson, R. (2006). Economic freedom of the world. Annual report, Fraser Institute.

Hofstede, G., & Bond, M. H. (1988). The Confucius connection: From cultural roots to economic growth. Organizational dynamics, 16(4), 5-21.
Ram, R. (2010). Social capital and happiness: Additional cross-country evidence. Journal of Happiness Studies, 11(4), 409-418.

Hagerty, M. R., & Veenhoven, R. (2003). Wealth and happiness revisited–growing national income does go with greater happiness. Social indicators research, 64(1), 1-27.

Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. Proceedings of the national academy of sciences, 107(38), 16489-16493.

Kenny, C. (1999). Does growth cause happiness, or does happiness cause growth?. Kyklos, 52(1), 3-25.

Ladaique, M. (2007). Measuring Well-being: What Role for Social Indicators?. Alternative Measures Of Well–Being. Society At A Glance: OECD Social Indicators. 19-36.

Layard, R. (2005). Rethinking public economics: The implications of rivalry and habit. Economics and happiness, 1(1), 147-170.

Lee, D. Y., Park, S. H., Uhlemann, M. R., & Patsult, P. (2000). What makes you happy?: A comparison of self-reported criteria of happiness between two cultures. Social Indicators Research, 50(3), 351-362.

Lu, L., Gilmour, R., & Kao, S. F. (2001). Cultural values and happiness: An East-West dialogue. The Journal of social psychology, 141(4), 477-493.

McBride, M. (2001). Relative-income effects on subjective well-being in the cross-section. Journal of Economic Behavior & Organization, 45(3), 251-278.

Mohanty, M. S. (2014). What determines happiness? Income or attitude: Evidence from the US longitudinal data. Journal of Neuroscience, Psychology, and Economics, 7(2), 80.

Oswald, A. J., Proto, E., & Sgroi, D. (2015). Happiness and productivity. Journal of Labor Economics, 33(4), 789-822.

Palmore, E., & Luikart, C. (1972). Health and social factors related to life satisfaction. Journal of Health and Social Behavior, 68-80.

Sacks, D. W., Stevenson, B., & Wolfers, J. (2012). The new stylized facts about income and subjective well-being. Emotion, 12(6), 1181.

Smith, S., & Razzell, P. (1975). The pools winners. Caliban Books.

Stevenson, B., & Wolfers, J. (2008). Economic growth and subjective well-being: Reassessing the Easterlin paradox (No. w14282). National Bureau of Economic Research.

Stiglitz, J. E., Sen, A., & Fitoussi, J. P. (2010). Report by the commission on the measurement of economic performance and social progress. Paris: Commission on the Measurement of Economic Performance and Social Progress.

Stutzer, A. (2004). The role of income aspirations in individual happiness. Journal of Economic Behavior & Organization, 54(1), 89-109.

The World Bank. World Bank Open Data. Retrieved May 10, 2017, from http://data.worldbank.org/

United Nations Development Programme, Human Development Report 2016. Retrieved May 25, 2017, from http://hdr.undp.org/en/2016-report

Veenhoven, R. (1988). The utility of happiness. Social indicators research, 20(4), 333-354.

Veenhoven, R. (1991). Is happiness relative?. Social indicators research, 24(1), 1-34.

Veenhoven, R. (2003). Hedonism and happiness. Journal of happiness studies, 4(4), 437-457.

Veenhoven, R. (2013). The four qualities of life ordering concepts and measures of the good life. In The exploration of happiness (pp. 195-226). Springer Netherlands.

White, M. D. (2014). The Problems with Measuring and Using Happiness for Policy Purposes. Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA.

8. Appendix

(i) Linear Indicator Models

VARIABLES	(1) (GNI)	(2) (NNI)	(3) (HHC)	(4) (HDI)
GNIht	3.077***			
NNIht		2.457***		
HHCons			6.233***	
HDI				5.973***
Unemployment Rate	-0.024***	-0.025***	-0.022***	-0.040***
Life Expectancy	0.052***	0.065***	0.047***	0.002
CO2 emissions per capita	0.004	0.011	0.015	0.001
Inflation Rate	0.002	-0.001	0.002	-0.004
Govt. Education Exp.	0.018***	0.021***	0.022***	0.031***
Time Fixed Effects	Yes	Yes	Yes	Yes
Constant	1.254**	0.496	1.398***	1.073
Observations	616	622	617	348
Number of Countries	118	123	119	115
Within R^2	0.0720	0.0538	0.0762	0.0608
Between R^2	0.7298	0.7317	0.7501	0.7094
Overall R^2	0.7010	0.7173	0.7144	0.6902

*** p<0.01, ** p<0.05, * p<0.1

(ii) Indicator & Happiness Correlation Table

2010	**Happ.**	**GDP**	**GNI**	**NNI**	**HHC**	**HDI**
Happ.	1					
GDP	0.7178	1				
GNI	0.7382	0.9906	1			
NNI	0.7458	0.9892	0.9993	1		
HHC	0.7403	0.9318	0.9492	0.9509	1	
HDI	0.7343	0.7034	0.7257	0.7252	0.7563	1
2011	**Happ.**	**GDP**	**GNI**	**NNI**	**HHC**	**HDI**
Happ.	1					
GDP	0.7111	1				
GNI	0.7403	0.9939	1			
NNI	0.7376	0.9848	0.9972	1		
HHC	0.7471	0.9127	0.9398	0.9347	1	
HDI	0.7386	0.6939	0.7318	0.7172	0.7552	1
2012	**Happ.**	**GDP**	**GNI**	**NNI**	**HHC**	**HDI**
Happ.	1					
GDP	0.7185	1				
GNI	0.7325	0.9914	1			
NNI	0.7443	0.9895	0.9945	1		
HHC	0.7458	0.9205	0.949	0.9335	1	
HDI	0.7665	0.6995	0.7272	0.7173	0.7588	1

2013	Happ.	GDP	GNI	NNI	HHC	HDI
Happ.	1					
GDP	0.7144	1				
GNI	0.7223	0.9921	1			
NNI	0.7456	0.9804	0.9928	1		
HHC	0.7321	0.951	0.9733	0.9644	1	
HDI	0.7816	0.7089	0.7438	0.744	0.7618	1
2014	**Happ.**	**GDP**	**GNI**	**NNI**	**HHC**	**HDI**
Happ.	1					
GDP	0.6832	1				
GNI	0.6959	0.9922	1			
NNI	0.709	0.9861	0.9933	1		
HHC	0.7149	0.9238	0.9515	0.9348	1	
HDI	0.8145	0.6927	0.7208	0.7209	0.7531	1
2015	**Happ.**	**GDP**	**GNI**	**NNI**	**HHC**	**HDI**
Happ.	1					
GDP	0.7439	1				
GNI	0.7457	0.9881	1			
NNI	0.7762	0.9884	0.9872	1		
HHC	0.7699	0.9604	0.9729	0.9724	1	
HDI	0.8197	0.7164	0.7317	0.743	0.7656	1

(iii) Hausmann Tests

Indicator	(1) (GDP)	(2) (GNI)	(3) (NNI)	(4) (HHC)	(5) (HDI)
Chi-Squared	-20.52	27.43	30.70	20.26	13.53
P-value	-	0.0109**	0.0037***	0.0623*	0.1956

*** p<0.01, ** p<0.05, * p<0.1

(iv) Regressions by Indicator & Region

1. GDP per capita

VARIABLES	(1) (East Asia & Pacific)	(2) (Europe & Central Asia)	(3) (Latin America & Caribbean)	(4) (Middle East & North Africa)	(6) (South Asia)	(7) (Sub-Saharan Africa)
GDP per capita ($100,000)	1.283*	0.853	8.360***	3.704**	-18.941	4.713
Unemployment Rate	-0.011	-0.037***	-0.030*	-0.069	0.008	-0.007
Life Expectancy	0.020	0.118***	0.018	0.084*	0.033	-0.004
CO2 emissions per capita	0.089**	0.073***	0.076	-0.060*	0.644	0.087***
Inflation Rate	-0.003	-0.006	-0.010	-0.010	0.006	0.017***
Govt. Education Exp.	0.059**	0.038	-0.006	0.010	0.018	0.016*
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.798	-3.856**	4.544	0.618	2.007	3.863***
Observations	81	250	98	35	39	127
Number of Countries	13	42	20	10	7	33
Within R^2	0.1611	0.2559	0.4512	0.3608	0.0559	0.3342
Between R^2	0.7478	0.7939	0.4096	0.6925	0.4625	0.2148
Overall R^2	0.7187	0.7934	0.3440	0.7542	0.2756	0.2956

2. GNI per capita

VARIABLES	(1) (East Asia & Pacific)	(2) (Europe & Central Asia)	(3) (Latin America & Caribbean)	(4) (Middle East & North Africa)	(6) (South Asia)	(7) (Sub-Saharan Africa)
GNI per capita ($100,000)	-0.050	2.685***	17.443***	7.231***	-12.407	8.024
Unemployment Rate	0.016	-0.029***	-0.043***	-0.011	0.070	-0.010
Life Expectancy	0.057	0.056*	0.009	0.072*	-0.036	-0.008
CO2 emissions per capita	0.103***	0.074***	-0.101	-0.077***	0.296	0.065
Inflation Rate	-0.007	-0.006	-0.014	0.014*	0.006	0.017***
Govt. Education Exp.	0.053*	0.019	0.017	-0.024	-0.004	0.015
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.079	0.585	4.656	0.522	6.835	4.107***
Observations	76	240	97	31	36	123
Number of Countries	12	39	19	7	7	32
Within R^2	0.1073	0.2961	0.4601	0.1774	0.0352	0.3456
Between R^2	0.7433	0.8489	0.5954	0.9729	0.2942	0.2308
Overall R^2	0.7113	0.8279	0.5176	0.9278	0.3928	0.2951

3. NNI per capita

VARIABLES	(1) (East Asia & Pacific)	(2) (Europe & Central Asia)	(3) (Latin America & Caribbean)	(4) (Middle East & North Africa)	(6) (South Asia)	(7) (Sub-Saharan Africa)
NNI per capita ($100,000)	1.301	1.860***	10.362***	2.939	-17.569	10.352
Unemployment Rate	0.002	-0.034***	-0.031*	-0.134	0.009	-0.009
Life Expectancy	0.056*	0.099***	0.019	0.170***	0.030	-0.007
CO2 emissions per capita	0.058*	0.067***	0.072	-0.057	0.613	0.065
Inflation Rate	-0.015	-0.005	-0.012	-0.009	0.008	0.015***
Govt. Education Exp.	0.084***	0.026	-0.007	0.042	0.019	0.017*
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.069	-2.372	4.470	-5.593*	2.141	4.010***
Observations	75	249	98	29	39	119
Number of Countries	12	42	20	8	7	32
Within R^2	0.1809	0.2623	0.4391	0.3810	0.0540	0.3305
Between R^2	0.8906	0.8174	0.4595	0.7694	0.5011	0.2365
Overall R^2	0.8468	0.8115	0.3957	0.8688	0.2688	0.2839

4. Household Consumption

VARIABLES	(1) (East Asia & Pacific)	(2) (Europe & Central Asia)	(3) (Latin America & Caribbean)	(4) (Middle East & North Africa)	(6) (South Asia)	(7) (Sub-Saharan Africa)
Household Consumption ($100,000)	0.122	4.895***	22.658***	25.979***	-7.541	22.094***
Unemployment Rate	0.015	-0.027**	-0.043***	-0.004	0.068	-0.007
Life Expectancy	0.054	0.069**	0.017	-0.170***	-0.047	-0.009
CO2 emissions per capita	0.102*	0.059**	-0.015	0.006	0.232	0.018
Inflation Rate	-0.006	-0.006	-0.011	0.004	0.009	0.017***
Govt. Education Exp.	0.054	0.043*	0.009	0.040***	-0.000	0.016
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.263	-0.607	4.142	16.113***	7.500	4.064***
Observations	76	243	97	31	36	121
Number of Countries	12	41	19	7	7	31
Within R^2	0.1092	0.2870	0.4462	0.2858	0.0361	0.3408
Between R^2	0.7433	0.8298	0.5323	0.9625	0.2998	0.3412
Overall R^2	0.7113	0.8226	0.4598	0.9475	0.3866	0.3411

5. Human Development Index

VARIABLES	(1) (East Asia & Pacific)	(2) (Europe & Central Asia)	(3) (Latin America & Caribbean)	(4) (Middle East & North Africa)	(5) (South Asia)	(6) (Sub-Saharan Africa)
Human Development Index (1-10)	5.301	5.274**	11.923***	9.430	-12.534***	3.189**
Unemployment Rate	-0.026	-0.052***	-0.024	0.003	0.140	-0.013
Life Expectancy	-0.005	0.105***	-0.009	-0.070	0.257***	-0.026
CO2 emissions per capita	0.034	0.027	-0.222	-0.004	1.188***	0.028
Inflation Rate	0.020	-0.006***	-0.062	-0.034	-0.016	0.009
Govt. Education Exp.	0.023	0.051*	0.017	-0.067	-0.019	0.028***
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	1.565	-6.794***	-0.890	5.486	-6.379	3.690***
Observations	42	129	51	18	24	78
Number of Countries	12	41	17	7	7	29
Within R^2	0.2961	0.1405	0.2915	0.1673	0.0915	0.5044
Between R^2	0.6877	0.8320	0.5257	0.6883	0.6963	0.2056
Overall R^2	0.7125	0.7852	0.3692	0.7780	0.5638	0.3371

(v) Pooled OLS Robustness Checks

1. Hypothesis 1

Variables	(1) (Only GDP)	(2) (Added Controls)	(3) (Full Model)	(4) (Excl. Health)	(5) (GDP-Squared)
GDP per capita ($100,000)	4.170***	0.503	0.707	1.821***	5.058***
GDP per capita^2	-	-	-	-	-3.659***
Unemployment Rate	-	-0.016***	-0.014**	-0.010	-0.015***
Life Expectancy	-	0.071***	0.071***	0.069***	0.056***
CO2 emissions per capita	-	0.013**	0.010*	-0.004	-0.004
Inflation Rate	-	0.003	0.003	0.004	0.008*
Govt. Education Exp.	-	0.045***	0.046***	0.037***	0.041***
Govt. Health Exp.	-	0.000***	0.000***	-	-
Time Fixed Effects	No	No	Yes	Yes	Yes
Constant	4.866***	-0.477	-0.413	-0.136	0.525
Observations	1,248	637	637	643	643
Number of Countries	158	126	126	127	127
R-squared	0.509	0.729	0.740	0.709	0.725

*** p<0.01, ** p<0.05, * p<0.1

2. Hypothesis 2

Variables	(1) (GNI)	(2) (NNI)	(3) (HHC)	(4) (HDI)
GNI per capita ($100,000)	5.438***			
GNI per capita^2	-3.685***			
NNI per capita ($100,000)		7.664***		
NNI per capita^2		-7.361***		
Household Cons. ($100,000)			11.057***	
Household Cons.^{2}			-15.699***	
Human Development Index (1-10)				-4.191**
Human Development Index^2				7.428***
Unemployment Rate	-0.011*	-0.019***	-0.017***	-0.039***
Life Expectancy	0.058***	0.057***	0.053***	0.012
CO2 emissions per capita	-0.006	-0.006	0.008**	-0.004
Inflation Rate	0.009**	0.007	0.009**	-0.010**
Govt. Education Exp.	0.041***	0.051***	0.045***	0.054***
Time Fixed Effects	Yes	Yes	Yes	Yes
Constant	0.319	0.341	0.533	3.339***
Observations	616	622	617	348
Number of Countries	118	123	119	115
R-squared	0.726	0.748	0.734	0.720

*** p<0.01, ** p<0.05, * p<0.1

3. Hypothesis 3

VARIABLES	(1) (East Asia & Pacific)	(2) (Europe & Central Asia)	(3) (Latin America & Caribbean)	(4) (Middle East & North Africa)	(5) (North America)	(6) (South Asia)	(7) (Sub-Saharan Africa)
GDP	1.283**	2.425***	14.224***	3.704**	-	-18.941	3.234
GNI	-0.050	3.074***	19.087***	7.231***	-	-12.407	5.550
NNI	1.301*	3.612***	18.645***	2.939	-	-17.569	10.524*
HHC	0.122	6.366***	28.188***	25.979***	-	-7.541	24.049***
HDI	2.310	6.171***	9.136**	9.430	-	-12.534***	2.806***

*** p<0.01, ** p<0.05, * p<0.1

(vi) Fixed Effects Robustness Checks

1. Hypothesis 1

Variables	(1) (Only GDP)	(2) (Added Controls)	(3) (Full Model)	(4) (Excl. Health)	(5) (GDP-Squared)
GDP per capita ($100,000)	1.362***	0.774	0.488	-0.092	1.296
GDP per capita^2	-	-	-	-	-1.174
Unemployment Rate	-	-0.029***	-0.033***	-0.034***	-0.031***
Life Expectancy	-	-0.006	-0.082***	-0.076***	-0.072***
CO2 emissions per capita	-	0.037	0.044*	0.050**	0.050**
Inflation Rate	-	-0.000	-0.001	-0.001	-0.000
Govt. Education Exp.	-	0.020***	0.022***	0.021***	0.020***
Govt. Health Exp.	-	0.000	-0.000	-	-
Time Fixed Effects	No	No	Yes	Yes	Yes
Constant	5.258***	5.612***	10.950***	10.580***	10.156***
Observations	1,248	637	637	643	643
Number of Countries	158	126	126	127	127
Within R^2	0.0125	0.0755	0.1106	0.1076	0.1096
Between R^2	0.5464	0.3087	0.4181	0.3633	0.2147
Overall R^2	0.5093	0.3144	0.3958	0.3390	0.1909

*** p<0.01, ** p<0.05, * p<0.1

2. Hypothesis 2

Variables	(1) (GNI)	(2) (NNI)	(3) (HHC)	(4) (HDI)
GNI per capita ($100,000)	12.407***			
GNI per capita^2	-10.457**			
NNI per capita ($100,000)		2.137		
NNI per capita^2		-2.355		
Household Cons. ($100,000)			11.178	
Household Cons.^{2}			-16.813	
Human Development Index (1-10)				-37.089***
Human Development Index^2				27.466***
Unemployment Rate	-0.024***	-0.032***	-0.026***	-0.035**
Life Expectancy	-0.055**	-0.067**	-0.058**	-0.009
CO2 emissions per capita	0.033	0.052**	0.042*	0.028
Inflation Rate	0.001	-0.002	0.001	-0.003
Govt. Education Exp.	0.015**	0.018**	0.018**	0.036***
Time Fixed Effects	Yes	Yes	Yes	Yes
Constant	7.898***	9.823***	8.674***	17.496***
Observations	616	622	617	348
Number of Countries	118	123	119	115
Within R^2	0.1223	0.1139	0.1126	0.1459
Between R^2	0.4388	0.1153	0.1587	0.0027
Overall R^2	0.4628	0.0928	0.2021	0.0065

*** p<0.01, ** p<0.05, * p<0.1

3. Hypothesis 3

VARIABLES	(1) (East Asia & Pacific)	(2) (Europe & Central Asia)	(3) (Latin America & Caribbean)	(4) (Middle East & North Africa)	(5) (North America)	(6) (South Asia)	(7) (Sub-Saharan Africa)
GDP	0 .097	-0.705	3.519	1.933	-	-11.747	19.362
GNI	-1.733	2.090	8.298	6.288	-	211.890	17.415
NNI	0 .111	-0.178	2.996	2.466	-	0 .111	22.285
HHC	0 .539	3.348	5.896	-41.068	-	270.998	103.626
HDI	7.596	8.284	14.060	45.566	-	-87.325	6.874

*** p<0.01, ** p<0.05, * p<0.1

(vii) Regressions by Income Group Robustness Checks

VARIABLES	(1) (High)	(2) (Upper-Middle)	(3) (Lower-Middle)	(4) (Low)
GDP per capita ($100,000)	6.541***	39.582***	67.361***	32.542
GDP per capita^2	-3.962***	-142.257**	-1,083.294**	-1,047.362
Unemployment Rate	-0.002	-0.047***	-0.044***	0.006
Life Expectancy	0.020	0.003	0.028***	0.007
CO2 emissions per capita	-0.011	-0.115***	-0.014	-0.068
Inflation Rate	-0.013	0.006	0.003	0.017***
Govt. Education Exp.	0.017	0.100***	0.048***	0.014
Time Fixed Effects	Yes	Yes	Yes	Yes
Constant	3.449**	3.214**	1.890***	3.032***
Observations	244	147	156	96
R-squared	0.615	0.624	0.439	0.132

*** p<0.01, ** p<0.05, * p<0.1

(viii) Dickey-Fuller Robustness Checks

REGION	Z(t)	p-value
All	-2.572	0.0989
East Asia & Pacific	-2.547	0.1045
Europe & Central Asia	-3.460	0.0091
Latin America & Caribbean	-2.506	0.1141
Middle East & North Africa	-3.862	0.0023
North America	-1.936	0.3154
South Asia	-2.668	0.0798
Sub-Saharan Africa	-2.551	0.1036

(ix) Interaction Variable Robustness Checks

VARIABLES	(1) (Sectors)	(2) (Sectors)	(3) (Sectors)	(4) (Sectors)	(5) (Sectors)
East Asia & Pacific	-	-	-	-	-
Europe & Central Asia	-0.388***	-0.340***	-0.395***	-0.370***	-3.130***
Latin America & Caribbean	0.472***	0.391***	0.351***	0.286**	-0.392
Middle East & North Africa	0.026	-0.062	-0.043	-0.668***	-2.390
North America	2.182***	6.296***	2.296***	2.708***	48.157***
South Asia	-0.359**	-0.254	-0.393**	-0.278	0.072
Sub-Saharan Africa	-0.974***	-1.095***	-1.008***	-1.137***	-3.651***
Indicator	6.505***	7.106***	8.810***	11.247***	-20.070***
Indicator^2	-5.670***	-6.587***	-9.617***	-16.494***	17.415***
Indicator*EAP	-	-	-	-	-
Indicator*ECA	1.991***	2.080***	2.026***	3.833***	4.075***
Indicator*LAC	5.954***	6.741***	7.635***	11.521***	1.918
Indicator*MENA	0.079	0.656	0.081	7.903***	3.173
Indicator*NA	-3.348**	-11.531***	-4.594***	-7.187***	-52.130***
Indicator*SA	1.339	-4.925	0.398	-4.855	-0.673
Indicator*SSA	4.816*	6.772**	11.074***	14.681***	5.311***
Unemployment Rate	-0.013**	-0.014***	-0.017***	-0.016***	-0.028***
Life Expectancy	-0.003	-0.008	0.001	-0.008	-0.012
CO2 emissions per capita	0.011	0.001	0.004	0.016***	0.007
Inflation Rate	0.010***	0.010**	0.009**	0.012***	-0.003
Govt. Education Exp.	0.032***	0.033***	0.037***	0.035***	0.048***
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Constant	4.724***	4.991***	4.455***	4.926***	10.731***
Observations	643	616	622	617	348
R-squared	0.833	0.835	0.844	0.839	0.813

*** p<0.01, ** p<0.05, * p<0.1